GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:22:17 ; Search time 0.731707 Seconds
(without alignments)
482.536 Million cell updates/sec

Title: US-09-890-752A-20
Perfect score: 61
Sequence: 1 PLSSIFSRIGDP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61	100.0	55	4	US-09-361-707-97	Sequence 97, Appl
2	61	100.0	174	3	US-08-480-173A-44	Sequence 44, Appl
3	61	100.0	174	3	US-08-484-408A-44	Sequence 44, Appl
4	61	100.0	281	1	US-08-105-483-214	Sequence 214, App
5	61	100.0	281	1	US-08-709-209-214	Sequence 214, App
6	61	100.0	281	1	US-08-458-101-214	Sequence 214, App
7	61	100.0	389	1	US-08-105-483-216	Sequence 216, App
8	61	100.0	389	1	US-08-105-483-219	Sequence 219, App
9	61	100.0	389	1	US-08-709-209-216	Sequence 216, App
10	61	100.0	389	1	US-08-709-209-219	Sequence 219, App
11	61	100.0	389	1	US-08-458-101-216	Sequence 216, App
12	61	100.0	389	1	US-08-458-101-219	Sequence 219, App
13	61	100.0	389	3	US-08-486-099-106	Sequence 106, App
14	61	100.0	389	3	US-08-360-107A-116	Sequence 116, App
15	61	100.0	389	3	US-08-484-223B-106	Sequence 106, App
16	61	100.0	389	3	US-08-919-597-106	Sequence 106, App
17	61	100.0	389	3	US-08-475-668A-106	Sequence 106, App
18	61	100.0	389	3	US-08-485-551A-106	Sequence 106, App
19	61	100.0	389	3	US-08-471-913A-106	Sequence 106, App
20	61	100.0	389	4	US-08-485-264A-106	Sequence 106, App
21	61	100.0	389	4	US-08-474-349A-106	Sequence 106, App
22	59	96.7	55	4	US-09-361-707-93	Sequence 93, Appl
23	59	96.7	55	4	US-09-361-707-95	Sequence 95, Appl
24	59	96.7	55	4	US-09-361-707-96	Sequence 96, Appl
25	59	96.7	174	2	US-08-683-262B-55	Sequence 55, Appl
26	59	96.7	174	3	US-08-480-173A-45	Sequence 45, Appl
27	59	96.7	174	3	US-08-484-408A-45	Sequence 45, Appl
28	59	96.7	174	4	US-09-361-707-55	Sequence 55, Appl
29	59	96.7	174	6	5204096-1	Patent No. 5204096
30	59	96.7	174	6	5204096-2	Patent No. 5204096
31	54	88.5	14	6	5164485-4	Patent No. 5164485
32	54	88.5	55	4	US-09-361-707-84	Sequence 84, Appl
33	54	88.5	55	4	US-09-361-707-85	Sequence 85, Appl
34	54	88.5	55	4	US-09-361-707-87	Sequence 87, Appl
35	54	88.5	55	4	US-09-361-707-89	Sequence 89, Appl
36	54	88.5	55	4	US-09-361-707-90	Sequence 90, Appl
37	54	88.5	170	2	US-08-683-262B-49	Sequence 49, Appl
38	54	88.5	170	4	US-09-361-707-49	Sequence 49, Appl
39	54	88.5	174	2	US-08-683-262B-46	Sequence 46, Appl
40	54	88.5	174	2	US-08-683-262B-48	Sequence 48, Appl
41	54	88.5	174	2	US-08-683-262B-56	Sequence 56, Appl
42	54	88.5	174	3	US-08-480-173A-48	Sequence 48, Appl
43	54	88.5	174	3	US-08-484-408A-48	Sequence 48, Appl
44	54	88.5	174	4	US-09-361-707-46	Sequence 46, Appl
45	54	88.5	174	4	US-09-361-707-48	Sequence 48, Appl

ALIGNMENTS

```
RESULT 1
US-09-361-707-97
; Sequence 97, Application US/09361707
; Patent No. 6258937
;  GENERAL INFORMATION:
;       APPLICANT: Tong, Shuping
;                  Li, Jisu
;                  Wands, Jack R.
;       TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
;       NUMBER OF SEQUENCES: 117
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fish & Richardson P.C.
;            STREET: 225 Franklin Street
;            CITY: Boston
;            STATE: MA
;            COUNTRY: USA
;            ZIP: 02110-2804
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/361,707
;            FILING DATE: 27-Jul-1999
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/683,262
;            FILING DATE: 18-JUL-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Creason, Gary L.
;            REGISTRATION NUMBER: 34,310
;            REFERENCE/DOCKET NUMBER: 00786/287003
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 542-5070
;            TELEFAX: (617) 542-8906
;  INFORMATION FOR SEQ ID NO: 97:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 55 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 97:
US-09-361-707-97

  Query Match           100.0%;  Score 61;  DB 4;  Length 55;
  Best Local Similarity  100.0%;  Pred. No. 9.2e-05;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db      41 PLSSIFSRIGDP 52
```

```
RESULT 2
US-08-480-173A-44
; Sequence 44, Application US/08480173A
; Patent No. 6072049
;  GENERAL INFORMATION:
;    APPLICANT:  Thoma, Hans A
;    TITLE OF INVENTION:  HEPATITIS B SURFACE ANTIGEN VACCINE
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Popovich & Wiles, P.A.
;      STREET:  80 S. 8th Street, Suite 1902
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,173A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Popovich, Thomas E
;      REGISTRATION NUMBER:  30,099
;      REFERENCE/DOCKET NUMBER:  MED1003USD4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-334-8991
;      TELEFAX:  612-334-8994
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  174 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-480-173A-44

  Query Match           100.0%;  Score 61;  DB 3;  Length 174;
  Best Local Similarity  100.0%;  Pred. No. 0.00034;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     160 PLSSIFSRIGDP 171


RESULT 3
US-08-484-408A-44
; Sequence 44, Application US/08484408A
; Patent No. 6117653
;  GENERAL INFORMATION:
;    APPLICANT:  Thoma, Hans A
;    TITLE OF INVENTION:  HEPATITIS B SURFACE ANTIGEN VACCINE
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Popovich & Wiles, P.A.
;      STREET:  80 S. 8th Street, Suite 1902
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,408A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Popovich, Thomas E
;      REGISTRATION NUMBER:  30,099
;      REFERENCE/DOCKET NUMBER:  MED1003USD4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-334-8991
;      TELEFAX:  612-334-8994
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  174 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-484-408A-44

  Query Match           100.0%;  Score 61;  DB 3;  Length 174;
  Best Local Similarity  100.0%;  Pred. No. 0.00034;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     160 PLSSIFSRIGDP 171


RESULT 4
US-08-105-483-214
; Sequence 214, Application US/08105483
; Patent No. 5494807
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/105,483
;      FILING DATE:  12-AUG-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  214:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  281 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-105-483-214

  Query Match           100.0%;  Score 61;  DB 1;  Length 281;
  Best Local Similarity  100.0%;  Pred. No. 0.0006;
```

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db     41 PLSSIFSRIGDP 52
```

RESULT 5
US-08-709-209-214
; Sequence 214, Application US/08709209
; Patent No. 5762938
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,209
; FILING DATE: 21-AUG-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/105,483
; FILING DATE: 12-AUG-1993
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 214:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-709-209-214

Query Match 100.0%; Score 61; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.0006;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db     41 PLSSIFSRIGDP 52
```

RESULT 6
US-08-458-101-214
; Sequence 214, Application US/08458101
; Patent No. 5766599
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Perkus, Marion E.
; APPLICANT: Taylor, Jill
; APPLICANT: Tartaglia, James
; APPLICANT: No. 5766599ton, Elizabeth K.
; APPLICANT: Riviere, Michel
; APPLICANT: de Taisne, Charles
; APPLICANT: Limbach, Keith J.
; APPLICANT: Johnson, Gerard P.
; APPLICANT: Pincus, Steven E.
; APPLICANT: Cox, William I.
; APPLICANT: Audonnet, Jean-Christophe Francis
; APPLICANT: Gettig, Russell Robert
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,101
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2740
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 214:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-458-101-214

Query Match 100.0%; Score 61; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.0006;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db     41 PLSSIFSRIGDP 52
```

RESULT 7
US-08-105-483-216
; Sequence 216, Application US/08105483
; Patent No. 5494807
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/105,483
;      FILING DATE:  12-AUG-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  216:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-105-483-216

  Query Match             100.0%;  Score 61;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.00086;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 8
US-08-105-483-219
; Sequence 219, Application US/08105483
; Patent No. 5494807
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Prommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/105,483
;      FILING DATE:  12-AUG-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-105-483-219

  Query Match             100.0%;  Score 61;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.00086;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 9
US-08-709-209-216
; Sequence 216, Application US/08709209
; Patent No. 5762938
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Prommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/709,209
;      FILING DATE:  21-AUG-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/105,483
;      FILING DATE:  12-AUG-1993
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  216:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-709-209-216

  Query Match             100.0%;  Score 61;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.00086;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 10
US-08-709-209-219
; Sequence 219, Application US/08709209
; Patent No. 5762938
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
```

```
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/709,209
;      FILING DATE:  21-AUG-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/105,483
;      FILING DATE:  12-AUG-1993
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-709-209-219

  Query Match             100.0%;  Score 61;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.00086;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 11
US-08-458-101-216
; Sequence 216, Application US/08458101
; Patent No. 5766599
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Perkus, Marion E.
;    APPLICANT:  Taylor, Jill
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  No. 5766599ton, Elizabeth K.
;    APPLICANT:  Riviere, Michel
;    APPLICANT:  de Taisne, Charles
;    APPLICANT:  Limbach, Keith J.
;    APPLICANT:  Johnson, Gerard P.
;    APPLICANT:  Pincus, Steven E.
;    APPLICANT:  Cox, William I.
;    APPLICANT:  Audonnet, Jean-Christophe Francis
;    APPLICANT:  Gettig, Russell Robert
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,101
;      FILING DATE:  01-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2740
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  216:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-458-101-216

  Query Match             100.0%;  Score 61;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.00086;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 12
US-08-458-101-219
; Sequence 219, Application US/08458101
; Patent No. 5766599
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Perkus, Marion E.
;    APPLICANT:  Taylor, Jill
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  No. 5766599ton, Elizabeth K.
;    APPLICANT:  Riviere, Michel
;    APPLICANT:  de Taisne, Charles
;    APPLICANT:  Limbach, Keith J.
;    APPLICANT:  Johnson, Gerard P.
;    APPLICANT:  Pincus, Steven E.
;    APPLICANT:  Cox, William I.
;    APPLICANT:  Audonnet, Jean-Christophe Francis
;    APPLICANT:  Gettig, Russell Robert
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,101
```

; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2740
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-458-101-219

Query Match 100.0%; Score 61; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 149 PLSSIFSRIGDP 160

RESULT 13
US-08-486-099-106
; Sequence 106, Application US/08486099
; Patent No. 6013263
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
; TITLE OF INVENTION: B VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,099
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-486-099-106

Query Match 100.0%; Score 61; DB 3; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 149 PLSSIFSRIGDP 160

RESULT 14
US-08-360-107A-116
; Sequence 116, Application US/08360107A
; Patent No. 6017536
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,107A
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-360-107A-116

Query Match 100.0%; Score 61; DB 3; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 149 PLSSIFSRIGDP 160

RESULT 15
US-08-484-223B-106
; Sequence 106, Application US/08484223B
; Patent No. 6020459

```
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  COMPOSITIONS FOR INHIBITION OF
;    TITLE OF INVENTION:  MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
;    TITLE OF INVENTION:  TRANSMISSION
;    NUMBER OF SEQUENCES:  245
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,223B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-029
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-484-223B-106

  Query Match           100.0%;  Score 61;  DB 3;  Length 389;
  Best Local Similarity  100.0%;  Pred. No. 0.00086;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


Search completed: February  3, 2003, 09:27:10
Job time : 0.731707 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:25:53 ; Search time 0.439024 Seconds
(without alignments)
551.547 Million cell updates/sec

Title: US-09-890-752A-20
Perfect score: 61
Sequence: 1 PLSSIFSRIGDP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61	100.0	281	9	US-09-247-890-12	Sequence 12, Appl
2	59	96.7	174	10	US-09-818-066-55	Sequence 55, Appl
3	59	96.7	389	9	US-09-821-877-2	Sequence 2, Appli
4	54	88.5	170	10	US-09-818-066-49	Sequence 49, Appl
5	54	88.5	174	10	US-09-818-066-46	Sequence 46, Appl
6	54	88.5	174	10	US-09-818-066-48	Sequence 48, Appl
7	54	88.5	174	10	US-09-818-066-56	Sequence 56, Appl
8	54	88.5	281	9	US-09-247-890-10	Sequence 10, Appl
9	54	88.5	400	9	US-10-209-264-3	Sequence 3, Appli
10	52	85.2	174	10	US-09-818-066-50	Sequence 50, Appl
11	52	85.2	174	10	US-09-818-066-58	Sequence 58, Appl
12	49	80.3	174	10	US-09-818-066-53	Sequence 53, Appl
13	46	75.4	174	10	US-09-818-066-45	Sequence 45, Appl
14	46	75.4	174	10	US-09-818-066-47	Sequence 47, Appl
15	44	72.1	174	10	US-09-818-066-37	Sequence 37, Appl
16	44	72.1	174	10	US-09-818-066-54	Sequence 54, Appl
17	43	70.5	174	10	US-09-818-066-57	Sequence 57, Appl
18	40	65.6	64	10	US-09-917-340-85	Sequence 85, Appl
19	40	65.6	174	10	US-09-818-066-42	Sequence 42, Appl
20	40	65.6	174	10	US-09-818-066-62	Sequence 62, Appl
21	38	62.3	174	10	US-09-818-066-38	Sequence 38, Appl
22	38	62.3	174	10	US-09-818-066-40	Sequence 40, Appl
23	38	62.3	174	10	US-09-818-066-41	Sequence 41, Appl
24	37	60.7	174	10	US-09-818-066-59	Sequence 59, Appl
25	36	59.0	55	10	US-09-879-257A-45	Sequence 45, Appl
26	36	59.0	174	10	US-09-818-066-35	Sequence 35, Appl
27	36	59.0	174	10	US-09-818-066-36	Sequence 36, Appl
28	36	59.0	174	10	US-09-818-066-52	Sequence 52, Appl
29	35	57.4	96	10	US-09-864-761-39796	Sequence 39796, A
30	35	57.4	173	9	US-10-051-643-207	Sequence 207, App
31	35	57.4	174	10	US-09-818-066-39	Sequence 39, Appl
32	35	57.4	690	10	US-09-815-242-12460	Sequence 12460, A
33	35	57.4	703	10	US-09-815-242-5610	Sequence 5610, Ap
34	35	57.4	705	9	US-09-738-626-4055	Sequence 4055, Ap
35	35	57.4	743	10	US-09-771-161A-164	Sequence 164, App
36	35	57.4	743	10	US-09-771-161A-254	Sequence 254, App
37	35	57.4	1036	10	US-09-771-161A-255	Sequence 255, App
38	35	57.4	1036	10	US-09-771-161A-256	Sequence 256, App
39	34	55.7	64	10	US-09-917-340-84	Sequence 84, Appl
40	34	55.7	119	10	US-09-925-299-799	Sequence 799, App
41	34	55.7	174	10	US-09-818-066-43	Sequence 43, Appl
42	34	55.7	174	10	US-09-818-066-44	Sequence 44, Appl
43	34	55.7	284	9	US-09-895-913A-102	Sequence 102, App
44	34	55.7	312	10	US-09-801-368-260	Sequence 260, App
45	34	55.7	478	9	US-10-156-634A-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-247-890-12
; Sequence 12, Application US/09247890
; Publication No. US20020198162A1
; GENERAL INFORMATION:
;  APPLICANT: Punnonen, Juha
;  APPLICANT: Bass, Steven H.
;  APPLICANT: Whalen, Robert Gerald
;  APPLICANT: Howard, Russell
;  APPLICANT: Stemmer, Willem P.C.
;  APPLICANT: Maxygen, Inc.
;  TITLE OF INVENTION: Antigen Library Immunization
;  FILE REFERENCE: 018097-028710US
;  CURRENT APPLICATION NUMBER: US/09/247,890
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: US 60/074,294
;  EARLIER FILING DATE: 1998-02-11
;  EARLIER APPLICATION NUMBER: US 60/105,509
;  EARLIER FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-247-890-12

  Query Match             100.0%;  Score 61;  DB 9;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.00052;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db      41 PLSSIFSRIGDP 52


RESULT 2
US-09-818-066-55
; Sequence 55, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;        APPLICANT: Shuping Tong et al.
```

```
;       TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;       NUMBER OF SEQUENCES: 75
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fish & Richardson P.C.
;            STREET: 225 Franklin Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02110-2804
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/818,066
;            FILING DATE: 27-Mar-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/683,262
;            FILING DATE: 18-JUL-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Fraser, Janis K.
;            REGISTRATION NUMBER: 31,819
;            REFERENCE/DOCKET NUMBER: 00786/287002
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 542-5070
;            TELEFAX: (617) 542-8906
;            TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 55:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 174 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-818-066-55

  Query Match             96.7%;  Score 59;  DB 10;  Length 174;
  Best Local Similarity   91.7%;  Pred. No. 0.00073;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:||||||||||
Db     160 PISSIFSRIGDP 171


RESULT 3
US-09-821-877-2
; Sequence 2, Application US/09821877
; Patent No. US20020177124A1
; GENERAL INFORMATION:
;  APPLICANT: Abbott Laboratories
;  APPLICANT:  Coleman, Paul F.
;  APPLICANT:  Mushahwar, Isa K.
;  TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant
;  TITLE OF INVENTION:  And Methods Of Detection Thereof
;  FILE REFERENCE: 6794.US.01
;  CURRENT APPLICATION NUMBER: US/09/821,877
;  CURRENT FILING DATE:  2000-03-03
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 389
;   TYPE: PRT
;   ORGANISM: Hepatitis B Virus
US-09-821-877-2

  Query Match             96.7%;  Score 59;  DB 9;  Length 389;
  Best Local Similarity   91.7%;  Pred. No. 0.0017;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:||||||||||
Db     149 PISSIFSRIGDP 160


RESULT 4
US-09-818-066-49
; Sequence 49, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;       APPLICANT: Shuping Tong et al.
;       TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;       NUMBER OF SEQUENCES: 75
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fish & Richardson P.C.
;            STREET: 225 Franklin Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02110-2804
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/818,066
;            FILING DATE: 27-Mar-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/683,262
;            FILING DATE: 18-JUL-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Fraser, Janis K.
;            REGISTRATION NUMBER: 31,819
;            REFERENCE/DOCKET NUMBER: 00786/287002
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 542-5070
;            TELEFAX: (617) 542-8906
;            TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 49:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 170 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-09-818-066-49

  Query Match             88.5%;  Score 54;  DB 10;  Length 170;
  Best Local Similarity   83.3%;  Pred. No. 0.0058;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:|||||| |||
Db     156 PISSIFSRTGDP 167


RESULT 5
US-09-818-066-46
; Sequence 46, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;       APPLICANT: Shuping Tong et al.
;       TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;       NUMBER OF SEQUENCES: 75
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fish & Richardson P.C.
;            STREET: 225 Franklin Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02110-2804
;       COMPUTER READABLE FORM:
```

```
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/818,066
;              FILING DATE: 27-Mar-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US 08/683,262
;              FILING DATE: 18-JUL-1996
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Fraser, Janis K.
;              REGISTRATION NUMBER: 31,819
;              REFERENCE/DOCKET NUMBER: 00786/287002
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (617) 542-5070
;              TELEFAX: (617) 542-8906
;              TELEX: 200154
;    INFORMATION FOR SEQ ID NO: 46:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 174 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-818-066-46

  Query Match             88.5%;  Score 54;  DB 10;  Length 174;
  Best Local Similarity   83.3%;  Pred. No. 0.006;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:|||||| |||
Db     160 PISSIFSRTGDP 171


RESULT 6
US-09-818-066-48
; Sequence 48, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;         APPLICANT: Shuping Tong et al.
;         TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;         NUMBER OF SEQUENCES: 75
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Fish & Richardson P.C.
;              STREET: 225 Franklin Street
;              CITY: Boston
;              STATE: Massachusetts
;              COUNTRY: U.S.A.
;              ZIP: 02110-2804
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/818,066
;              FILING DATE: 27-Mar-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US 08/683,262
;              FILING DATE: 18-JUL-1996
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Fraser, Janis K.
;              REGISTRATION NUMBER: 31,819
;              REFERENCE/DOCKET NUMBER: 00786/287002
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (617) 542-5070
;              TELEFAX: (617) 542-8906
;              TELEX: 200154
;    INFORMATION FOR SEQ ID NO: 48:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 174 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-09-818-066-48

  Query Match             88.5%;  Score 54;  DB 10;  Length 174;
  Best Local Similarity   83.3%;  Pred. No. 0.006;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:|||||| |||
Db     160 PISSIFSRTGDP 171


RESULT 7
US-09-818-066-56
; Sequence 56, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;         APPLICANT: Shuping Tong et al.
;         TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;         NUMBER OF SEQUENCES: 75
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Fish & Richardson P.C.
;              STREET: 225 Franklin Street
;              CITY: Boston
;              STATE: Massachusetts
;              COUNTRY: U.S.A.
;              ZIP: 02110-2804
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/818,066
;              FILING DATE: 27-Mar-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US 08/683,262
;              FILING DATE: 18-JUL-1996
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Fraser, Janis K.
;              REGISTRATION NUMBER: 31,819
;              REFERENCE/DOCKET NUMBER: 00786/287002
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (617) 542-5070
;              TELEFAX: (617) 542-8906
;              TELEX: 200154
;    INFORMATION FOR SEQ ID NO: 56:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 174 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-09-818-066-56

  Query Match             88.5%;  Score 54;  DB 10;  Length 174;
  Best Local Similarity   100.0%;  Pred. No. 0.006;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGD 11
           |||||||||||
Db     160 PLSSIFSRIGD 170


RESULT 8
US-09-247-890-10
; Sequence 10, Application US/09247890
```

```
; Publication No. US20020198162A1
; GENERAL INFORMATION:
;  APPLICANT: Punnonen, Juha
;  APPLICANT: Bass, Steven H.
;  APPLICANT: Whalen, Robert Gerald
;  APPLICANT: Howard, Russell
;  APPLICANT: Stemmer, Willem P.C.
;  APPLICANT: Maxygen, Inc.
;  TITLE OF INVENTION: Antigen Library Immunization
;  FILE REFERENCE: 018097-028710US
;  CURRENT APPLICATION NUMBER: US/09/247,890
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: US 60/074,294
;  EARLIER FILING DATE: 1998-02-11
;  EARLIER APPLICATION NUMBER: US 60/105,509
;  EARLIER FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-247-890-10

  Query Match             88.5%;  Score 54;  DB 9;  Length 281;
  Best Local Similarity   83.3%;  Pred. No. 0.01;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:|||||| |||
Db      41 PISSIFSRTGDP 52


RESULT 9
US-10-209-264-3
; Sequence 3, Application US/10209264
; Publication No. US20030003111A1
;   GENERAL INFORMATION:
;        APPLICANT: Oon, Chong Jin
;                   Lim, Gek Keow
;                   Zhao, Yi
;                   Chen, Wei Ning
;        TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
;                            USES THEREOF
;        NUMBER OF SEQUENCES: 11
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Ladas & Parry
;             STREET: 26 West 61 Street
;             CITY: New York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10023
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/209,264
;             FILING DATE: 31-Jul-2002
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/SG98/00046
;             FILING DATE: 19-JAN-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mass, Clifford J.
;             REGISTRATION NUMBER: 30,086
;             REFERENCE/DOCKET NUMBER: U-013109-7
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 708-1800
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 400 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-209-264-3

  Query Match             88.5%;  Score 54;  DB 9;  Length 400;
  Best Local Similarity   83.3%;  Pred. No. 0.015;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:|||||| |||
Db     160 PISSIFSRTGDP 171


RESULT 10
US-09-818-066-50
; Sequence 50, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;        APPLICANT: Shuping Tong et al.
;        TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;        NUMBER OF SEQUENCES: 75
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/818,066
;             FILING DATE: 27-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/683,262
;             FILING DATE: 18-JUL-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Fraser, Janis K.
;             REGISTRATION NUMBER: 31,819
;             REFERENCE/DOCKET NUMBER: 00786/287002
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 542-5070
;             TELEFAX: (617) 542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 50:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 174 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-09-818-066-50

  Query Match             85.2%;  Score 52;  DB 10;  Length 174;
  Best Local Similarity   90.9%;  Pred. No. 0.014;
  Matches   10;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LSSIFSRIGDP 12
           :||||||||||
Db     161 ISSIFSRIGDP 171


RESULT 11
US-09-818-066-58
; Sequence 58, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
```

```
;         APPLICANT: Shuping Tong et al.
;         TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;         NUMBER OF SEQUENCES: 75
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Fish & Richardson P.C.
;              STREET: 225 Franklin Street
;              CITY: Boston
;              STATE: Massachusetts
;              COUNTRY: U.S.A.
;              ZIP: 02110-2804
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/818,066
;              FILING DATE: 27-Mar-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US 08/683,262
;              FILING DATE: 18-JUL-1996
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Fraser, Janis K.
;              REGISTRATION NUMBER: 31,819
;              REFERENCE/DOCKET NUMBER: 00786/287002
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (617) 542-5070
;              TELEFAX: (617) 542-8906
;              TELEX: 200154
;    INFORMATION FOR SEQ ID NO: 58:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 174 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-09-818-066-58

  Query Match             85.2%;  Score 52;  DB 10;  Length 174;
  Best Local Similarity   90.9%;  Pred. No. 0.014;
  Matches   10;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LSSIFSRIGDP 12
           :||||||||||
Db     161 ISSIFSRIGDP 171


RESULT 12
US-09-818-066-53
; Sequence 53, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;         APPLICANT: Shuping Tong et al.
;         TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;         NUMBER OF SEQUENCES: 75
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Fish & Richardson P.C.
;              STREET: 225 Franklin Street
;              CITY: Boston
;              STATE: Massachusetts
;              COUNTRY: U.S.A.
;              ZIP: 02110-2804
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/818,066
;              FILING DATE: 27-Mar-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US 08/683,262
;              FILING DATE: 18-JUL-1996
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Fraser, Janis K.
;              REGISTRATION NUMBER: 31,819
;              REFERENCE/DOCKET NUMBER: 00786/287002
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (617) 542-5070
;              TELEFAX: (617) 542-8906
;              TELEX: 200154
;    INFORMATION FOR SEQ ID NO: 53:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 174 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-09-818-066-53

  Query Match             80.3%;  Score 49;  DB 10;  Length 174;
  Best Local Similarity   81.8%;  Pred. No. 0.049;
  Matches    9;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LSSIFSRIGDP 12
           :||||:|||||
Db     161 ISSIFTRIGDP 171


RESULT 13
US-09-818-066-45
; Sequence 45, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;         APPLICANT: Shuping Tong et al.
;         TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;         NUMBER OF SEQUENCES: 75
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Fish & Richardson P.C.
;              STREET: 225 Franklin Street
;              CITY: Boston
;              STATE: Massachusetts
;              COUNTRY: U.S.A.
;              ZIP: 02110-2804
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/818,066
;              FILING DATE: 27-Mar-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US 08/683,262
;              FILING DATE: 18-JUL-1996
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Fraser, Janis K.
;              REGISTRATION NUMBER: 31,819
;              REFERENCE/DOCKET NUMBER: 00786/287002
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (617) 542-5070
;              TELEFAX: (617) 542-8906
;              TELEX: 200154
;    INFORMATION FOR SEQ ID NO: 45:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 174 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-818-066-45

  Query Match             75.4%;  Score 46;  DB 10;  Length 174;
```

```
  Best Local Similarity   75.0%;  Pred. No. 0.18;
  Matches    9;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:||| || |||
Db     160 PISSISSRTGDP 171


RESULT 14
US-09-818-066-47
; Sequence 47, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;        APPLICANT: Shuping Tong et al.
;        TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;        NUMBER OF SEQUENCES: 75
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/818,066
;             FILING DATE: 27-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/683,262
;             FILING DATE: 18-JUL-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Fraser, Janis K.
;             REGISTRATION NUMBER: 31,819
;             REFERENCE/DOCKET NUMBER: 00786/287002
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 542-5070
;             TELEFAX: (617) 542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 47:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 174 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-818-066-47

  Query Match             75.4%;  Score 46;  DB 10;  Length 174;
  Best Local Similarity   75.0%;  Pred. No. 0.18;
  Matches    9;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:||| || |||
Db     160 PISSISSRTGDP 171


RESULT 15
US-09-818-066-37
; Sequence 37, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;        APPLICANT: Shuping Tong et al.
;        TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;        NUMBER OF SEQUENCES: 75
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/818,066
;             FILING DATE: 27-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/683,262
;             FILING DATE: 18-JUL-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Fraser, Janis K.
;             REGISTRATION NUMBER: 31,819
;             REFERENCE/DOCKET NUMBER: 00786/287002
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 542-5070
;             TELEFAX: (617) 542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 37:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 174 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-09-818-066-37

  Query Match             72.1%;  Score 44;  DB 10;  Length 174;
  Best Local Similarity   81.8%;  Pred. No. 0.41;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LSSIFSRIGDP 12
           :||| ||||||
Db     161 ISSISSRIGDP 171


Search completed: February  3, 2003, 09:35:29
Job time : 0.439024 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:24:27 ; Search time 0.982578 Seconds
(without alignments)
988.252 Million cell updates/sec

Title: US-09-890-752A-20
Perfect score: 61
Sequence: 1 PLSSIFSRIGDP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 404799 seqs, 80919614 residues

Total number of hits satisfying chosen parameters: 404799

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61	100.0	12	6	US-10-211-088-302	Sequence 302, App
2	59	96.7	64	6	US-10-326-908-19	Sequence 19, Appl
3	59	96.7	67	6	US-10-326-908-21	Sequence 21, Appl
4	54	88.5	400	1	PCT-US02-32263-46	Sequence 46, Appl
5	54	88.5	400	6	US-10-287-994-46	Sequence 46, Appl
6	48	78.7	10	1	PCT-US02-13609-23	Sequence 23, Appl
7	39	63.9	399	1	PCT-US02-40225-3144	Sequence 3144, Ap
8	39	63.9	399	6	US-10-320-797-3144	Sequence 3144, Ap
9	37	60.7	533	6	US-10-092-411A-4053	Sequence 4053, Ap
10	35	57.4	53	1	PCT-US02-32727-2247	Sequence 2247, Ap
11	35	57.4	53	6	US-10-057-498-2247	Sequence 2247, Ap
12	35	57.4	70	1	PCT-US02-32727-19192	Sequence 19192, A
13	35	57.4	70	6	US-10-057-498-19192	Sequence 19192, A
14	35	57.4	144	1	PCT-US02-32727-23079	Sequence 23079, A
15	35	57.4	144	6	US-10-057-498-23079	Sequence 23079, A
16	35	57.4	199	6	US-10-218-140-3758	Sequence 3758, Ap
17	35	57.4	244	5	US-09-724-676-70237	Sequence 70237, A
18	35	57.4	244	5	US-09-724-676-70238	Sequence 70238, A
19	35	57.4	244	5	US-09-724-676A-70237	Sequence 70237, A
20	35	57.4	244	5	US-09-724-676A-70238	Sequence 70238, A
21	35	57.4	309	1	PCT-US02-34321-5	Sequence 5, Appli
22	35	57.4	318	1	PCT-US02-34321-3	Sequence 3, Appli
23	35	57.4	322	1	PCT-US02-32727-29962	Sequence 29962, A
24	35	57.4	687	5	US-09-724-676-70230	Sequence 70230, A
25	35	57.4	687	5	US-09-724-676-70231	Sequence 70231, A
26	35	57.4	687	5	US-09-724-676A-70230	Sequence 70230, A
27	35	57.4	687	5	US-09-724-676A-70231	Sequence 70231, A
28	35	57.4	696	5	US-09-724-676-70232	Sequence 70232, A
29	35	57.4	696	5	US-09-724-676-70233	Sequence 70233, A
30	35	57.4	696	5	US-09-724-676A-70232	Sequence 70232, A
31	35	57.4	696	5	US-09-724-676A-70233	Sequence 70233, A
32	35	57.4	741	1	PCT-US02-32727-1231	Sequence 1231, Ap
33	35	57.4	741	6	US-10-057-498-1231	Sequence 1231, Ap
34	35	57.4	1037	5	US-09-724-676-70223	Sequence 70223, A
35	35	57.4	1037	5	US-09-724-676A-70223	Sequence 70223, A
36	35	57.4	1047	5	US-09-724-676-70234	Sequence 70234, A
37	35	57.4	1047	5	US-09-724-676-70239	Sequence 70239, A
38	35	57.4	1047	5	US-09-724-676A-70234	Sequence 70234, A
39	35	57.4	1047	5	US-09-724-676A-70239	Sequence 70239, A
40	35	57.4	1069	5	US-09-724-676-70240	Sequence 70240, A
41	35	57.4	1069	5	US-09-724-676-70241	Sequence 70241, A
42	35	57.4	1069	5	US-09-724-676A-70240	Sequence 70240, A
43	35	57.4	1069	5	US-09-724-676A-70241	Sequence 70241, A
44	35	57.4	1075	5	US-09-724-676-70224	Sequence 70224, A
45	35	57.4	1075	5	US-09-724-676-70225	Sequence 70225, A

ALIGNMENTS

RESULT 1
US-10-211-088-302
; Sequence 302, Application US/10211088
; GENERAL INFORMATION:
; APPLICANT: Bright, Gary R.
; APPLICANT: Premkumar, D. David
; APPLICANT: Chen, Yih-Tai
; TITLE OF INVENTION: Novel Fusion Proteins And Assays For Molecular Binding
; FILE REFERENCE: 01-1022-US
; CURRENT APPLICATION NUMBER: US/10/211,088
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/309,395
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/341,589
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 302
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Protein-derived transport peptide
US-10-211-088-302

Query Match 100.0%; Score 61; DB 6; Length 12;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 1 PLSSIFSRIGDP 12

RESULT 2
US-10-326-908-19
; Sequence 19, Application US/10326908
; GENERAL INFORMATION:
; APPLICANT: Christopher JONES
; APPLICANT: Andrew BACON
; APPLICANT: Gill DOUCE
; APPLICANT: Mark PAGE
; TITLE OF INVENTION: Designing Immunogens
; FILE REFERENCE: 117-433 / N78027D
; CURRENT APPLICATION NUMBER: US/10/326,908
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 10/110,036
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB00/03857

; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: GB 0007789.1
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: GB 9923902.2
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: MS Word
; SEQ ID NO 19
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide immunogen
US-10-326-908-19

Query Match 96.7%; Score 59; DB 6; Length 64;
Best Local Similarity 91.7%; Pred. No. 0.00061;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
|:||||||||||||
Db 48 PISSIFSRIGDP 59

RESULT 3
US-10-326-908-21
; Sequence 21, Application US/10326908
; GENERAL INFORMATION:
; APPLICANT: Christopher JONES
; APPLICANT: Andrew BACON
; APPLICANT: Gill DOUCE
; APPLICANT: Mark PAGE
; TITLE OF INVENTION: Designing Immunogens
; FILE REFERENCE: 117-433 / N78027D
; CURRENT APPLICATION NUMBER: US/10/326,908
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 10/110,036
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB00/03857
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: GB 0007789.1
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: GB 9923902.2
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: MS Word
; SEQ ID NO 21
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide immunogen
US-10-326-908-21

Query Match 96.7%; Score 59; DB 6; Length 67;
Best Local Similarity 91.7%; Pred. No. 0.00064;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
|:||||||||||||
Db 48 PISSIFSRIGDP 59

RESULT 4
PCT-US02-32263-46
; Sequence 46, Application PC/TUS0232263
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bowe, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5050WO
; CURRENT APPLICATION NUMBER: PCT/US02/32263
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/334,233
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/334,301
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Hepatitis B virus
PCT-US02-32263-46

Query Match 88.5%; Score 54; DB 1; Length 400;
Best Local Similarity 83.3%; Pred. No. 0.036;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
|:||||| |||
Db 160 PISSIFSRTGDP 171

RESULT 5
US-10-287-994-46
; Sequence 46, Application US/10287994
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bowe, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5052-00
; CURRENT APPLICATION NUMBER: US/10/287,994
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 400

; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-287-994-46

Query Match 88.5%; Score 54; DB 6; Length 400;
Best Local Similarity 83.3%; Pred. No. 0.036;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
:|:|||||| |||
Db 160 PISSIFSRTGDP 171

RESULT 6
PCT-US02-13609-23
; Sequence 23, Application PC/TUS0213609
; GENERAL INFORMATION:
; APPLICANT: Targeted Genetics Corporation
; APPLICANT: Emerald Gene Systems, LTD
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 226272005340
; CURRENT APPLICATION NUMBER: PCT/US02/13609
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
PCT-US02-13609-23

Query Match 78.7%; Score 48; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIG 10
||||||||||
Db 1 PLSSIFSRIG 10

RESULT 7
PCT-US02-40225-3144
; Sequence 3144, Application PC/TUS0240225
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-228
; CURRENT APPLICATION NUMBER: PCT/US02/40225
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3144
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
PCT-US02-40225-3144

Query Match 63.9%; Score 39; DB 1; Length 399;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSIFSRIGD 11
||::|||||:
Db 213 LSAVFSRIGE 222

RESULT 8
US-10-320-797-3144
; Sequence 3144, Application US/10320797
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3144
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
US-10-320-797-3144

Query Match 63.9%; Score 39; DB 6; Length 399;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSIFSRIGD 11
||::|||||:
Db 213 LSAVFSRIGE 222

RESULT 9
US-10-092-411A-4053
; Sequence 4053, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 4053
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-4053

Query Match 60.7%; Score 37; DB 6; Length 533;
Best Local Similarity 60.0%; Pred. No. 82;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIG 10
||:||::|:|
Db 98 PLTSIYNRLG 107

RESULT 10
PCT-US02-32727-2247
; Sequence 2247, Application PC/TUS0232727
; GENERAL INFORMATION:

```
;  APPLICANT: Mitcham, Jennifer
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Persing, David
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Maisonneuve, Jean Francois
;  APPLICANT:  Zhang, Yanni
;  APPLICANT:  Wang, Siqing
;  APPLICANT:  Jen, Shyian
;  APPLICANT:  Lodes, Michael
;  APPLICANT:  Benson, Darin
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Barth, Brenda
;  APPLICANT:  Douglass,John
;  TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes V
;  FILE REFERENCE: 210121.514C1
;  CURRENT APPLICATION NUMBER: PCT/US02/32727
;  CURRENT FILING DATE:  2002-10-11
;  NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 2247
;   LENGTH: 53
;   TYPE: PRT
;   ORGANISM: Propioni acnes
PCT-US02-32727-2247

  Query Match             57.4%;  Score 35;  DB 1;  Length 53;
  Best Local Similarity   70.0%;  Pred. No. 18;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 SSIFSRIGDP 12
           ||: |||| |
Db      14 SSLHSRIGSP 23


RESULT 11
US-10-057-498-2247
; Sequence 2247, Application US/10057498
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Persing, David
;  TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes V
;  FILE REFERENCE: 210121.514
;  CURRENT APPLICATION NUMBER: US/10/057,498
;  CURRENT FILING DATE:  2001-04-20
;  NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 2247
;   LENGTH: 53
;   TYPE: PRT
;   ORGANISM: Propioni acnes
US-10-057-498-2247

  Query Match             57.4%;  Score 35;  DB 6;  Length 53;
  Best Local Similarity   70.0%;  Pred. No. 18;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 SSIFSRIGDP 12
           ||: |||| |
Db      14 SSLHSRIGSP 23


RESULT 12
PCT-US02-32727-19192
; Sequence 19192, Application PC/TUS0232727
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Persing, David
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Maisonneuve, Jean Francois
;  APPLICANT:  Zhang, Yanni
;  APPLICANT:  Wang, Siqing
;  APPLICANT:  Jen, Shyian
;  APPLICANT:  Lodes, Michael
;  APPLICANT:  Benson, Darin
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Barth, Brenda
;  APPLICANT:  Douglass,John
;  TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes Vi
;  FILE REFERENCE: 210121.514C1
;  CURRENT APPLICATION NUMBER: PCT/US02/32727
;  CURRENT FILING DATE:  2002-10-11
;  NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 19192
;   LENGTH: 70
;   TYPE: PRT
;   ORGANISM: Propioni acnes
PCT-US02-32727-19192

  Query Match             57.4%;  Score 35;  DB 1;  Length 70;
  Best Local Similarity   77.8%;  Pred. No. 24;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 SIFSRIGDP 12
           ||||| | |
Db       7 SIFSRFGKP 15


RESULT 13
US-10-057-498-19192
; Sequence 19192, Application US/10057498
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Persing, David
;  TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes Vi
;  FILE REFERENCE: 210121.514
;  CURRENT APPLICATION NUMBER: US/10/057,498
;  CURRENT FILING DATE:  2001-04-20
;  NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 19192
;   LENGTH: 70
;   TYPE: PRT
;   ORGANISM: Propioni acnes
US-10-057-498-19192

  Query Match             57.4%;  Score 35;  DB 6;  Length 70;
  Best Local Similarity   77.8%;  Pred. No. 24;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 SIFSRIGDP 12
           ||||| | |
Db       7 SIFSRFGKP 15


RESULT 14
PCT-US02-32727-23079
; Sequence 23079, Application PC/TUS0232727
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Persing, David
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Maisonneuve, Jean Francois
;  APPLICANT:  Zhang, Yanni
;  APPLICANT:  Wang, Siqing
;  APPLICANT:  Jen, Shyian
;  APPLICANT:  Lodes, Michael
;  APPLICANT:  Benson, Darin
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Barth, Brenda
;  APPLICANT:  Douglass,John
;  TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes Vi
;  FILE REFERENCE: 210121.514C1
```

; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 23079
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Propioni acnes
PCT-US02-32727-23079

Query Match 57.4%; Score 35; DB 1; Length 144;
Best Local Similarity 45.5%; Pred. No. 51;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LSSIFSRIGDP 12
|:::| :|||
Db 24 LAAVFKSLGDP 34

RESULT 15
US-10-057-498-23079
; Sequence 23079, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes V
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 23079
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Propioni acnes
US-10-057-498-23079

Query Match 57.4%; Score 35; DB 6; Length 144;
Best Local Similarity 45.5%; Pred. No. 51;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LSSIFSRIGDP 12
|:::| :|||
Db 24 LAAVFKSLGDP 34

Search completed: February 3, 2003, 09:35:01
Job time : 1.98258 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:57 ; Search time 0.794425 Seconds
(without alignments)
1452.136 Million cell updates/sec

Title: US-09-890-752A-20
Perfect score: 61
Sequence: 1 PLSSIFSRIGDP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61	100.0	378	2	S41869	surface antigen -
2	61	100.0	389	1	SAVLAH	large surface anti
3	61	100.0	389	2	S20749	surface antigen -
4	59	96.7	389	1	SAVLBH	large surface anti
5	59	96.7	389	1	SAVLAJ	large surface anti
6	59	96.7	389	1	SAVLA1	large surface anti
7	59	96.7	389	2	S47407	surface antigen -
8	59	96.7	389	2	S32202	large surface anti
9	54	88.5	382	2	T13474	large surface anti
10	54	88.5	384	2	T13469	large surface anti
11	54	88.5	389	2	S20745	surface antigen -
12	54	88.5	400	1	SAVLA	large surface anti
13	54	88.5	400	2	S35528	surface antigen -
14	54	88.5	445	2	S43492	surface antigen -
15	47	77.0	389	2	S67506	large surface anti
16	45	73.8	378	2	S41870	surface antigen -
17	45	73.8	389	2	S20753	surface antigen -
18	43	70.5	358	2	B81194	conserved hypothet
19	43	70.5	358	2	A81831	hypothetical prote
20	41	67.2	502	2	C87709	ubiquinone biosynt
21	40	65.6	389	1	SAVLCP	large surface anti
22	40	65.6	389	2	S41871	surface antigen -
23	39	63.9	399	2	AF2349	hypothetical prote
24	39	63.9	557	4	S42226	hypothetical large
25	38	62.3	278	2	AF2895	hypothetical prote
26	38	62.3	315	2	A86807	hypothetical prote
27	38	62.3	389	1	SAVLJ1	large surface anti
28	38	62.3	389	1	SAVLJ2	large surface anti
29	38	62.3	449	2	F86852	amino acid permeas
30	38	62.3	537	2	S67434	probable sterol O-
31	38	62.3	603	2	H69121	hypothetical prote
32	38	62.3	1786	1	H71527	probable excinucle
33	37	60.7	134	2	A82455	hypothetical prote
34	37	60.7	134	2	G82477	hypothetical prote
35	37	60.7	229	2	A97370	hypothetical prote
36	37	60.7	229	2	AI2587	conserved hypothet
37	37	60.7	333	2	A90956	probable integrase
38	37	60.7	333	2	F85804	integrase for prop
39	37	60.7	422	2	I50516	retinoid X recepto
40	37	60.7	495	2	T21115	hypothetical prote
41	37	60.7	533	1	S74401	hypothetical prote
42	37	60.7	582	2	D84362	hypothetical prote
43	37	60.7	1037	2	E84723	hypothetical prote
44	36.5	59.8	335	2	G86326	protein F18O14.7 [
45	36	59.0	230	2	A75260	hypothetical prote

ALIGNMENTS

RESULT 1
S41869
surface antigen - hepatitis B virus (subtype ayw, patient A) (fragment)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (midd
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient A
C;Date: 06-Jan-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
C;Accession: S41869
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.
submitted to the EMBL Data Library, January 1994
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg chronic
A;Reference number: S41869
A;Accession: S41869
A;Molecule type: DNA
A;Residues: 1-378 <LAI>
A;Cross-references: EMBL:X77309; NID:g452611; PIDN:CAA54515.1; PID:g452612
A;Experimental source: subtype ayw, patient A
C;Genetics:
A;Gene: S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-378/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-108/Domain: pre-S1 domain #status predicted <PRE1>
F;109-378/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PS
F;109-163/Domain: pre-S2 domain #status predicted <PRE2>
F;164-378/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 100.0%; Score 61; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160
```

RESULT 2
SAVLAH
large surface antigen - hepatitis B virus (subtype ayw)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 07-May-1999
C;Accession: A03703; JQ2064; PQ0591
R;Galibert, F.; Mandart, E.; Fitoussi, F.; Tiollais, P.; Charnay, P.
Nature 281, 646-650, 1979
A;Title: Nucleotide sequence of the hepatitis B virus genome (subtype ayw) in E. coli.
A;Reference number: A93214; MUID:81012091; PMID:399327
A;Accession: A03703
A;Molecule type: DNA
A;Residues: 1-389 <GAL>
A;Cross-references: GB:J02203

```
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Accession: JQ2064
A;Molecule type: DNA
A;Residues: 164-389 <NOR>
A;Experimental source: genogroup D, subtype ayw3
R;Norder, H.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 3141-3145, 1992
A;Title: Molecular basis of hepatitis B virus serotype variations within the four major
A;Reference number: PQ0453; MUID:93107848; PMID:1469353
A;Accession: PQ0591
A;Molecule type: DNA
A;Residues: 264-343 <NO2>
A;Experimental source: subtype ayw3, Kli
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             100.0%;  Score 61;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.0012;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 3
S20749
surface antigen - hepatitis B virus (subtype ayw, patient CI)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen S (small en
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient CI
C;Date: 20-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
C;Accession: S20749
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submitted to the EMBL Data Library, March 1992
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati
A;Reference number: S20745
A;Accession: S20749
A;Molecule type: DNA
A;Residues: 1-389 <LAI>
A;Cross-references: EMBL:X65258; NID:g59434; PIDN:CAA46353.1; PID:g59435
A;Experimental source: subtype ayw, patient CI
A;Note: typical methionine for the beginning of pre-S2 domain at position 109 is missing
C;Genetics:
A;Gene: S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

  Query Match             100.0%;  Score 61;  DB 2;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.0012;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 4
SAVLBH
large surface antigen - hepatitis B virus (subtype ayw, strain pHB320)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
A;Note: host Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 28-Jul-2000
C;Accession: A03704; PQ0585
R;Bichko, V.; Pushko, P.; Dreilina, D.; Pumpen, P.; Gren, E.
FEBS Lett. 185, 208-212, 1985
A;Title: Subtype ayw variant of hepatitis B virus: DNA primary structure analysis.
A;Reference number: A05237; MUID:85204397; PMID:3996597
A;Accession: A03704
A;Molecule type: DNA
A;Residues: 1-389 <BIC>
A;Cross-references: GB:X02496; NID:g62280; PIDN:CAB41701.1; PID:g4704321
R;Norder, H.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 3141-3145, 1992
A;Title: Molecular basis of hepatitis B virus serotype variations within the four major s
A;Reference number: PQ0453; MUID:93107848; PMID:1469353
A;Accession: PQ0585
A;Molecule type: DNA
A;Residues: 264-343 <NOR>
A;Experimental source: subtype ayw2, Tav
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             96.7%;  Score 59;  DB 1;  Length 389;
  Best Local Similarity   91.7%;  Pred. No. 0.0028;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:||||||||||
Db     149 PISSIFSRIGDP 160


RESULT 5
SAVLAJ
large surface antigen - hepatitis B virus (subtype adyw)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Nov-1996
C;Accession: A93217; A03703
R;Pasek, M.; Goto, T.; Gilbert, W.; Zink, B.; Schaller, H.; MacKay, P.; Leadbetter, G.; M
Nature 282, 575-579, 1979
A;Title: Hepatitis B virus genes and their expression in E. coli.
A;Reference number: A93217; MUID:81012115; PMID:399329
A;Accession: A93217
A;Molecule type: DNA
A;Residues: 1-389 <PAS>
A;Cross-references: GB:J02202
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             96.7%;  Score 59;  DB 1;  Length 389;
  Best Local Similarity   91.7%;  Pred. No. 0.0028;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:||||||||||
Db     149 PISSIFSRIGDP 160


RESULT 6
SAVLA1
large surface antigen - hepatitis B virus (strain alpha1)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
```

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C;Accession: B34773; JQ2082
R;Tong, S.; Li, J.; Vitvitski, L.; Trepo, C.
Virology 176, 596-603, 1990
A;Title: Active hepatitis B virus replication in the presence of anti-HBe is associated
A;Reference number: A34773; MUID:90266476; PMID:2345966
A;Accession: B34773
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-389 <TON>
A;Cross-references: EMBL:M32138; NID:g329667; PIDN:AAA45502.1; PID:g329669
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Accession: JQ2082
A;Molecule type: DNA
A;Residues: 164-389 <NOR>
A;Experimental source: genogroup D, subtype ayw2, strain HBV-alpha
C;Genetics:
A;Gene: pre-S1/pre-S2/S
A;Introns: 111/3
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.7%; Score 59; DB 1; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.0028;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
|:||||||||||
Db 149 PISSIFSRIGDP 160

RESULT 7
S47407
surface antigen - hepatitis B virus (subtype ayw4)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw4
C;Date: 23-Nov-1994 #sequence_revision 12-May-1995 #text_change 26-Aug-1999
C;Accession: S47407; JQ2071
R;Pluciennicczak, A.
submitted to the EMBL Data Library, August 1994
A;Description: Molecular cloning and sequencing of two complete genomes of polish isolat
A;Reference number: S47404
A;Accession: S47407
A;Molecule type: DNA
A;Residues: 1-389 <PLU>
A;Cross-references: EMBL:Z35716; NID:g527435; PIDN:CAA84788.1; PID:g527439
A;Experimental source: subtype ayw4
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Contents: genogroup D
A;Accession: JQ2071
A;Molecule type: DNA
A;Residues: 164-389 <NOR>
A;Experimental source: subtype ayw3, strain 8950/90
C;Genetics:
A;Gene: S
A;Introns: 111/3
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-108/Domain: pre-S1 domain #status predicted <PRE1>
F;109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
F;109-163/Domain: pre-S2 domain #status predicted <PRE2>
F;164-389/Product: surface antigen S (small envelope protein) #status predicted <SAG>

Query Match 96.7%; Score 59; DB 2; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.0028;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
|:||||||||||
Db 149 PISSIFSRIGDP 160

RESULT 8
S32202
large surface antigen - hepatitis B virus (subtype ayw, isolate patient C1005)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, isolate patient C1005
C;Date: 20-Feb-1995 #sequence_revision 06-Dec-1996 #text_change 06-Dec-1996
C;Accession: S32202
R;Preisler-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Rasenack, J.
submitted to the EMBL Data Library, March 1993
A;Description: Identification and sequence analysis of hepatitis B virus DNA in immunolog
A;Reference number: S32202
A;Accession: S32202
A;Molecule type: DNA
A;Residues: 1-389 <PRE>
A;Cross-references: EMBL:X72702
A;Experimental source: subtype ayw, isolate patient C1005
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-389/Product: large surface antigen (gene pre-S1/pre-S2/S) #status predicted <DSL>
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>

Query Match 96.7%; Score 59; DB 2; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.0028;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
|:||||||||||
Db 149 PISSIFSRIGDP 160

RESULT 9
T13474
large surface antigen - hepatitis B virus (isolate 29Y11HCC)
N;Alternate names: envelope protein; HBs antigen
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate 29Y11HCC
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
C;Accession: T13474
R;Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.
Arch. Virol. 143, 2313-2326, 1998
A;Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinon
A;Reference number: Z17684; MUID:99129050; PMID:9930189
A;Accession: T13474
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-382 <TAK>
A;Cross-references: EMBL:AB014388; NID:g3582381; PIDN:BAA32936.1; PID:g3582384
A;Experimental source: Japanese patient with hepatocellular carcinoma isolate 29Y11HCC
C;Genetics:
A;Gene: S
A;Introns: 123/2
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen

Query Match 88.5%; Score 54; DB 2; Length 382;
Best Local Similarity 83.3%; Pred. No. 0.025;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 PLSSIFSRIGDP 12
          |:|||||| |||
Db    142 PISSIFSRTGDP 153
```

RESULT 10
T13469
large surface antigen - hepatitis B virus (isolate 09D09HCC)
N;Alternate names: envelope protein
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate 09D09HCC
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
C;Accession: T13469
R;Takahashi, K.; Akahane,.Y.; Hino, K.; Ohta, Y.; Mishiro, S.
Arch. Virol. 143, 2313-2326, 1998
A;Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcino
A;Reference number: Z17684; MUID:99129050; PMID:9930189
A;Accession: T13469
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-384 <TAK>
A;Cross-references: EMBL:AB014368; NID:g3551304; PIDN:BAA32864.1; PID:g3551308
A;Experimental source: Japanese patient with hepatocellular carcinoma isolate 09D09HCC
C;Genetics:
A;Gene: S
A;Introns: 123/2
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen

```
  Query Match             88.5%;  Score 54;  DB 2;  Length 384;
  Best Local Similarity   83.3%;  Pred. No. 0.025;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          |:|||||| |||
Db    144 PISSIFSRTGDP 155
```

RESULT 11
S20745
surface antigen - hepatitis B virus (subtype ayw, patient C)
N;Alternate names: HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient C
C;Date: 20-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
C;Accession: S20745
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submitted to the EMBL Data Library, March 1992
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati
A;Reference number: S20745
A;Accession: S20745
A;Molecule type: DNA
A;Residues: 1-389 <LAI>
A;Cross-references: EMBL:X65257; NID:g59429; PIDN:CAA46349.1; PID:g59430
A;Experimental source: subtype ayw, patient C
C;Genetics:
A;Gene: S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-108/Domain: pre-S1 domain #status predicted <PRE1>
F;109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
F;109-163/Domain: pre-S2 domain #status predicted <PRE2>
F;164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

```
  Query Match             88.5%;  Score 54;  DB 2;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.025;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LSSIFSRIGDP 12
          ||||||||||
Db    150 LSSIFSRIGDP 160
```

RESULT 12
SAVLA
large surface antigen - hepatitis B virus (subtype adr)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 07-May-1999
C;Accession: A03705; S04569; JQ2107; PQ0608
R;Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Sugino, Y.; Nishioka, K.
Nucleic Acids Res. 11, 1747-1757, 1983
A;Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype a
A;Reference number: A93460; MUID:83168919; PMID:6300776
A;Accession: A03705
A;Molecule type: DNA
A;Residues: 1-400 <ONO>
A;Cross-references: GB:V00867
R;Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.
Nucleic Acids Res. 17, 2124, 1989
A;Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtype
A;Reference number: S04568; MUID:89183619; PMID:2928116
A;Accession: S04569
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-50,'R',52-66,'YP',69-129,'V',131-142,'P',144-164,'S',166-176,'S',178-338,'C
A;Cross-references: EMBL:X14193
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Accession: JQ2107
A;Molecule type: DNA
A;Residues: 175-400 <NOR>
A;Experimental source: genogroup C, subtype adr, strain pBRHBadr4
R;Norder, H.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 3141-3145, 1992
A;Title: Molecular basis of hepatitis B virus serotype variations within the four major s
A;Reference number: PQ0453; MUID:93107848; PMID:1469353
A;Accession: PQ0608
A;Molecule type: DNA
A;Residues: 275-354 <NO2>
A;Experimental source: subtype adrq+, Bau
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;120-400/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;175-400/Product: major surface antigen (gene S) #status predicted <MSA>
F;15,123,177/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
  Query Match             88.5%;  Score 54;  DB 1;  Length 400;
  Best Local Similarity   83.3%;  Pred. No. 0.026;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          |:|||||| |||
Db    160 PISSIFSRTGDP 171
```

RESULT 13
S35528
surface antigen - hepatitis B virus (subtype adr)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype adr
C;Date: 09-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C;Accession: S35528
R;Mukaide, M.; Kumazawa, T.; Hoshi, A.; Kawaguchi, R.; Hikiji, K.
Nucleic Acids Res. 20, 6105, 1992
A;Title: The complete nucleotide sequence of hepatitis B virus, subtype adr (SRADR) and p

```
A;Reference number: S35527; MUID:93096607; PMID:1461746
A;Accession: S35528
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-400 <MUK>
A;Cross-references: EMBL:D12980; NID:g221500; PIDN:BAA02359.1; PID:g221502
A;Experimental source: subtype adr
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992
C;Genetics:
A;Gene: S
A;Introns: 165/2
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-400/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-119/Domain: pre-S1 domain #status predicted <PRE1>
F;120-400/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
F;120-174/Domain: pre-S2 domain #status predicted <PRE2>
F;175-400/Product: surface antigen S (small envelope protein) #status predicted <PSD>

  Query Match             88.5%;  Score 54;  DB 2;  Length 400;
  Best Local Similarity   83.3%;  Pred. No. 0.026;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:|||||| |||
Db     160 PISSIFSRTGDP 171


RESULT 14
S43492
surface antigen - hepatitis B virus (subtype adr)
N;Alternate names: HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype adr
C;Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C;Accession: S43492; PN0601; PN0602
R;Loncarevic, I.F.; Zentgraf, H.; Schroeder, C.H.
Nucleic Acids Res. 18, 4940, 1990
A;Title: Sequence of a replication competent hepatitis B virus genome with a preX open r
A;Reference number: S12598; MUID:90370503; PMID:2395664
A;Accession: S43492
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-445 <LON>
A;Cross-references: EMBL:X52939; NID:g457780; PIDN:CAA37114.1; PID:g457783
A;Experimental source: subtype adr
R;Mimms, L.T.; Solomon, L.R.; Ebert, J.W.; Fields, H.
Biochem. Biophys. Res. Commun. 195, 186-191, 1993
A;Title: Unique PreS sequence in a gibbon-derived hepatitis B virus variant.
A;Reference number: PN0601; MUID:93371402; PMID:8363598
A;Accession: PN0601
A;Molecule type: protein
A;Residues: 57-95 <MIM>
A;Accession: PN0602
A;Molecule type: protein
A;Residues: 165-196 <MI2>
C;Genetics:
A;Gene: S
A;Introns: 210/3
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-445/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-164/Domain: pre-S1 domain #status predicted <PRE1>
F;165-445/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
F;165-219/Domain: pre-S2 domain #status predicted <PRE2>
F;220-445/Product: surface antigen S (small envelope protein) #status predicted <PSD>

  Query Match             88.5%;  Score 54;  DB 2;  Length 445;
  Best Local Similarity   83.3%;  Pred. No. 0.029;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:|||||| |||
Db     205 PISSIFSRTGDP 216


RESULT 15
S67506
large surface antigen - hepatitis B virus (subtype ayw3, isolate Hope CH1357)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw3, isolate Hope CH1357
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 26-Aug-1999
C;Accession: S67506
R;Norder, H.; Ebert, J.W.; Fields, H.A.; Mushahwar, I.K.; Magnius, L.O.
Virology 218, 214-223, 1996
A;Title: Complete sequencing of a gibbon hepatitis B virus genome reveals a unique genot
A;Reference number: S67503; MUID:96207410; PMID:8615024
A;Accession: S67506
A;Molecule type: DNA
A;Residues: 1-389 <NOR>
A;Cross-references: EMBL:U46935; NID:g1814218; PIDN:AAB41952.1; PID:g1814222
A;Experimental source: subtype ayw3, isolate Hope CH1357
C;Genetics:
A;Gene: pre-S1/pre-S2/S
A;Introns: 112/2
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-108/Domain: pre-S1-specific domain #status predicted <PRE1>
F;109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
F;109-163/Domain: pre-S2-specific domain #status predicted <PRE2>
F;164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>
F;4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             77.0%;  Score 47;  DB 2;  Length 389;
  Best Local Similarity   81.8%;  Pred. No. 0.54;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LSSIFSRIGDP 12
           :|||||| |||
Db     150 ISSIFSRTGDP 160


Search completed: February  3, 2003, 09:26:29
Job time : 1.79443 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:15:41 ; Search time 1.92335 Seconds
(without alignments)
831.368 Million cell updates/sec

Title: US-09-890-752A-20
Perfect score: 61
Sequence: 1 PLSSIFSRIGDP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61	100.0	12	21	AAB10598	HBV fusion protein
2	61	100.0	12	21	AAY87906	HBV cell surface p
3	61	100.0	25	12	AAR14747	HBV Pre S2 peptide
4	61	100.0	36	23	AAE17023	Hepatitis B virus
5	61	100.0	60	16	AAR76747	HBV surface antige
6	61	100.0	102	20	AAY30947	Human HbsAG (PreS2
7	61	100.0	164	21	AAY93740	Amino acid sequenc
8	61	100.0	164	21	AAY90760	Hepatitis B virus
9	61	100.0	174	12	AAR15617	HBsAg pre-S region
10	61	100.0	174	13	AAR23867	Pre-S gene region
11	61	100.0	215	21	AAB10597	HBV fusion protein
12	61	100.0	281	13	AAR27471	spsAg protein. Sy
13	61	100.0	281	20	AAY32835	HBsAg PreS2-S regi
14	61	100.0	389	13	AAR27472	lpsAg protein. Sy
15	61	100.0	389	13	AAR27474	I3L promoter/S12/c
16	61	100.0	393	22	AAU14097	Hepatitis B virus
17	59	96.7	55	8	AAP70002	Sequence encoded b
18	59	96.7	64	22	AAB81133	Hepatitis B virus
19	59	96.7	64	23	AAE17027	Hepatitis B virus
20	59	96.7	67	22	AAB81135	Hepatitis B virus
21	59	96.7	67	23	AAE17029	C-terminally trunc
22	59	96.7	174	12	AAR15618	HBsAg pre-S region
23	59	96.7	174	13	AAR23868	Pre-S gene region
24	59	96.7	347	21	AAB10596	HBV fusion protein
25	59	96.7	389	22	AAG66931	HBV genotype D pre
26	54	88.5	174	12	AAR12394	Pre-S antigen. He
27	54	88.5	174	12	AAR15622	HBsAg pre-S region
28	54	88.5	174	13	AAR23871	Pre-S gene region
29	54	88.5	281	7	AAP60163	Subtype adr hepati
30	54	88.5	281	7	AAP60560	Hepatitis B virus
31	54	88.5	281	7	AAP60617	C-terminal defecti
32	54	88.5	281	8	AAP70294	Subtype adr HBsAg
33	54	88.5	281	15	AAR62870	Hepatitis B Virus
34	54	88.5	281	20	AAY32834	HBsAg PreS2-S regi
35	54	88.5	281	22	AAM48419	Protein #3 used to
36	54	88.5	389	7	AAP60794	Adr-type hepatitis
37	54	88.5	400	17	AAR93801	Hepatitis B virus
38	54	88.5	400	21	AAY54045	Amino acid seqeunc
39	52	85.2	54	15	AAR47945	Recombinant transf
40	52	85.2	55	22	AAE04710	Hepatitis B virus
41	52	85.2	174	10	AAP91483	Translation produc
42	52	85.2	389	22	AAG62931	Amino acid sequenc
43	52	85.2	399	22	AAG66919	HBV HBPol protein.
44	52	85.2	399	22	AAG66932	HBV genotype E pre
45	51	83.6	175	10	AAP91482	Translation produc

ALIGNMENTS

RESULT 1
AAB10598
ID AAB10598 standard; Protein; 12 AA.
XX
AC AAB10598;
XX
DT 08-JAN-2001 (first entry)
XX
DE HBV fusion protein cell permeability peptide fragment.
XX
KW Fusion protein; protein coat; virus-specific packaging signal; psi;
KW virus protein; cell permeability; cell-specific binding site; LHB;
KW large surface protein; core antigen; gene therapy.
XX
OS Hepatitis b virus.
XX
PN WO200046376-A2.
XX
PD 10-AUG-2000.
XX
PF 04-FEB-2000; 2000WO-DE00363.
XX
PR 05-FEB-1999; 99DE-1004800.
XX
PA (HILD/) HILDT E.
XX
PI Hildt E, Hofschneider P;
XX
DR WPI; 2000-514959/46.
DR N-PSDB; AAA71751.
XX
PT Particle for cell-specific gene delivery, useful in gene therapy,
PT comprises nucleic acid in protein coat that includes a fusion protein

PT of viral protein, permeability peptide and cell-binding site -
XX
PS Disclosure; Page 2; 34pp; German.
XX
CC This invention describes a novel particle (A), comprising a protein coat
CC with a fusion protein (FP), and, inside the coat, a nucleic acid (I)
CC including the sequence for a virus-specific packaging signal (psi) and a
CC structural gene. FP contains a virus protein (VP), a peptide (P) that
CC mediates cell permeability and a heterologous cell-specific binding site
CC (RGD). The invention also describes (1) producing (A) in which FP
CC contains an LHBs (large surface protein of hepatitis B virus (HBV)) and
CC RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),
CC (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector
CC containing the DNA of (d). The products of the invention are used in gene
CC therapy of cells and tissues, in vivo or ex vivo. This sequence
CC represents a fragment of the Hepatitis b virus cell permeablilty peptide
CC which is described in the method of the invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 61; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 1 PLSSIFSRIGDP 12

RESULT 2
AAY87906
ID AAY87906 standard; Protein; 12 AA.
XX
AC AAY87906;
XX
DT 01-SEP-2000 (first entry)
XX
DE HBV cell surface protein ZPP fragment.
XX
KW Cell surface protein; ZPP; cell permeability mediating polypeptide;
KW therapy; treatment.
XX
OS Hepatitis B virus.
XX
PN DE19850718-C1.
XX
PD 18-MAY-2000.
XX
PF 03-NOV-1998; 98DE-1050718.
XX
PR 03-NOV-1998; 98DE-1050718.
XX
PA (HILD/) HILDT E.
XX
PI Hildt E, Oess S;
XX
DR WPI; 2000-340689/30.
DR N-PSDB; AAA39621.
XX
PT Novel cell permeability-mediating polypeptide useful for mediating
PT permeability of substances such as other polypeptides in cells -
XX
PS Claim 1; Fig 1; 8pp; German.
XX
CC This invention describes a novel cell permeability mediating polypeptide
CC (CPM) comprising (A) or a sequence which differs by one or more amino
CC acids, which shows cell permeability and is not hepatitis B virus (HBV)
CC surface protein and where the DNA sequence of the latter amino acid
CC sequence hybridizes with the DNA of (I). CPM can be used to mediate cell
CC permeability of substances, especially polypeptides, nucleic acids and
CC chemical connections. This is useful for increasing the reactivity of
CC such substances in therapeutic treatment of various conditions. This
CC sequence represents a fragment of the HBV cell surface protein ZPP which
CC is described in the method of the invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 61; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 1 PLSSIFSRIGDP 12

RESULT 3
AAR14747
ID AAR14747 standard; peptide; 25 AA.
XX
AC AAR14747;
XX
DT 28-JAN-1992 (first entry)
XX
DE HBV Pre S2 peptide.
XX
KW Monoclonal antibodies; hepatitis B virus; envelope; vaccine;
KW epitope; antigen.
XX
OS Synthetic.
XX
PN EP456215-A.
XX
PD 13-NOV-1991.
XX
PF 08-MAY-1991; 91EP-0107488.
XX
PR 11-MAY-1990; 90US-0522286.
XX
PA (ABBO) ABBOTT LABORATORIES.
XX
PI Mimms LT, Floreani MF;
XX
DR WPI; 1991-334198/46.
XX
PT New MAbs to PreS2 and PreS1 polypeptide(s) of Hepatitis-B
PT envelope - used to sub-type HBV in sample and as vaccine against
PT HBV.
XX
PS Claim 1; Page 22; 24pp; English.
XX
CC The peptide corresponds to residues 150-174 of the HBV PreS2 protein.
CC It was used to raise monoclonal antibodies specific for the HBV M
CC protein but which do not bind to the HBV S or L proteins. The MAb
CC can be used to sub-type HBV and to prepare subunit vcaccines for HBV.
XX
SQ Sequence 25 AA;

Query Match 100.0%; Score 61; DB 12; Length 25;
Best Local Similarity 100.0%; Pred. No. 8.3e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 11 PLSSIFSRIGDP 22

RESULT 4
AAE17023
ID AAE17023 standard; peptide; 36 AA.
XX
AC AAE17023;
XX
DT 18-APR-2002 (first entry)
XX
DE Hepatitis B virus (HBV) ayw subtype pre-S2 region.

```
XX
KW   Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
KW   prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
KW   hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
KW   tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
KW   dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
KW   food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
KW   antiprotozoal; pre-S2 region.
XX
OS   Hepatitis B virus.
XX
PN   WO200198333-A2.
XX
PD   27-DEC-2001.
XX
PF   22-JUN-2001; 2001WO-GB02817.
XX
PR   22-JUN-2000; 2000GB-0015308.
PR   06-OCT-2000; 2000GB-0024544.
XX
PA   (CELL-) CELLTECH PHARM LTD.
XX
PI   Page M,  Li J,  Pumpens P;
XX
DR   WPI; 2002-098223/13.
DR   N-PSDB; AAD27424.
XX
PT   New proteins comprising a modified hepatitis B core antigen, useful as
PT   a vaccine in prophylactic or therapeutic vaccination of the human or
PT   animal body, particularly against hepatitis B virus infection -
XX
PS   Disclosure; Fig 3A; 40pp; English.
XX
CC   The invention relates to modified proteins comprising hepatitis B virus
CC   (HBV) core antigen (HBcAg) wherein one or more of the four arginine
CC   repeats has been deleted and the protein comprising the C-terminal
CC   cysteine of HBcAg. The deleted region may be replaced by an epitope
CC   from a protein other than HBcAg, in which case the HBcAg acts as a
CC   carrier to present the epitope to the immune system. This chimeric
CC   protein or its nucleic acid is useful as a vaccine or in a method of
CC   prophylactic or therapeutic vaccination of the human or animal body,
CC   particularly against HBV. The nucleic acid encoding the protein may
CC   be used in gene therapy or DNA vaccination protocols. The chimeric
CC   protein or its nucleic acid may also be used as the basis of a
CC   prophylactic vaccine against a range of diseases, e.g. HBV, hepatitis
CC   A virus (HAV), hepatitis C virus (HCV), influenza, foot-and-mouth
CC   disease, polio, herpes, rabies, acquired immunodeficiency syndrome
CC   (AIDS), dengue fever, yellow fever, malaria, tuberculosis, whooping
CC   cough, salmonellosis, typhoid, food poisoning, diarrhoea, meningitis
CC   or gonorrhea. The present sequence is Hepatitis B virus (HBV) ayw
CC   subtype pre-S2 region.
XX
SQ   Sequence   36 AA;

  Query Match             100.0%;  Score 61;  DB 23;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 0.00012;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db      22 PLSSIFSRIGDP 33


RESULT 5
AAR76747
ID   AAR76747 standard; Peptide; 60 AA.
XX
AC   AAR76747;
XX
DT   18-MAR-1996  (first entry)
XX
DE   HBV surface antigen pre-S2 region and FimH residues 224-226.
XX
KW   FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
KW   FimA; FimF; FimG; receptor binding site; PCR; amplify.
XX
OS   Chimeric - Hepatitis B virus.
OS   Chimeric - Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Represents FimH residue 224"
FT   Peptide         2..4
FT                   /note= "Linker peptide"
FT   Peptide         5..56
FT                   /note= "HBV surface antigen pre-S2 region"
FT   Peptide         57..59
FT                   /note= "Linker peptide"
FT   Misc-difference 60
FT                   /note= "Represents FimH residue 226"
XX
PN   WO9520657-A1.
XX
PD   03-AUG-1995.
XX
PF   27-JAN-1995;   95WO-DK00042.
XX
PR   27-JAN-1994;   94US-0187166.
XX
PA   (GXBI-) GX BIOSYSTEMS AS.
XX
PI   Hasty DL,  Klemm P,  Molin S,  Pallesen L,  Sokurenko EV;
XX
DR   WPI; 1995-275442/36.
DR   N-PSDB; AAQ93059.
XX
PT   Receptor specific bacterial adhesins - useful for targetting active
PT   compounds and microbial cells to locations of receptors
XX
PS   Example 3; Page 57; 152pp; English.
XX
CC   This sequence is encoded by a fragment of the plasmid pLPA38, which
CC   comprises a fimH fusion gene comprising the sequence encoding the pre-S2
CC   region of the hepatitis B viral surface antigen inserted into the fimH
CC   gene.  The primer sequences given in AAQ93057-58 caused the insertion of
CC   the pre-S2 region into the FimH protein at position 224-226.  The
CC   resulting chimeric gene was then further modified by insertion of the
CC   cholera toxin B chain into a different position of the FimH adhesin of
CC   type 1 fimbriae.  Restiction site handles (BglII-sites) were introduced
CC   into the fimH gene, and the foreign epitopes are then inserted in-frame.
CC   In the selected positions the insertion of the epitopes did not
CC   significantly alter the adhesive function of the FimH protein.  The
CC   expression of the chimeric proteins on the surface of fimbriae on
CC   bacterial hosts illustrated the possibility of using bacterial adhesins
CC   as general presenters of foreign antigens and epitopes. These chimeric
CC   genes may be used in the production of variant FimH adhesins which may
CC   be useful for targetting active compounds and microbial cells to
CC   locations comprising selected receptors to which the adhesins bind.
XX
SQ   Sequence   60 AA;

  Query Match             100.0%;  Score 61;  DB 16;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 0.00021;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db      42 PLSSIFSRIGDP 53


RESULT 6
AAY30947
ID   AAY30947 standard; Protein; 102 AA.
XX
AC   AAY30947;
XX
```

DT 20-OCT-1999 (first entry)
XX
DE Human HbsAG (PreS2 plus S region) truncated peptide.
XX
KW Vaccine; accessory molecule; antigen; transport; presentation; allergy;
KW antibacterial; antiviral; antifungal; anti-allergic; antidiabetic;
KW anti-inflammatory; anti-arthritic; anti-asthma; anticancer; treatment;
KW prevention; infection; bacterial; viral; fungal; autoimmune disease;
KW rheumatoid arthritis; diabetes; multiple sclerosis; pancreatitis;
KW inflammatory condition; psoriasis; immune deficiency; metastases;
KW cancer; Th1-type helper response; humoral; cellular response; B cell;
KW T cell; proliferation; immunogolobulin synthesis; isotype switching;
KW immunomodulator; immune response; asthma; human; HbsAG.
XX
OS Homo sapiens.
XX
PN WO9941368-A2.
XX
PD 19-AUG-1999.
XX
PF 10-FEB-1999; 99WO-US03020.
XX
PR 11-FEB-1998; 98US-0074294.
PR 11-FEB-1998; 98US-0021769.
XX
PA (MAXY-) MAXYGEN INC.
XX
PI Howard R, Punnonen J, Stemmer WPC, Whalen RG;
XX
DR WPI; 1999-508645/42.
XX
PT Identifying nucleic acid that directly or indirectly modulates the
PT immune response to a genetic vaccine vector, e.g. for prevention of
PT infection or cancer
XX
PS Example 3; Page 70; 105pp; English.
XX
CC This invention describes the identification of a novel polynucleotide (I)
CC that modulates the immune response to a genetic vaccine vector (A), or
CC encodes a polypeptide (II) with similar effect. The invention also
CC describes the identification of a polynucleotide (Ia) encoding an
CC accessory molecule (IIa) that improves transport and presentation of
CC antigen by a cell. The products of the invention have antibacterial,
CC antiviral, antifungal, anti-allergic, antidiabetic, anti-inflammatory,
CC anti-arthritic, anti-asthma, anticancer and immunomodulatory. Optimized
CC (I) are incorporated into (A), or (I) or its encoded (II) are
CC administered together with (A). (A) are used to treat or prevent
CC infections (bacterial, viral or fungal), autoimmune disease (e.g.
CC rheumatoid arthritis, diabetes or multiple sclerosis), other inflammatory
CC conditions (e.g. psoriasis or pancreatitis), immune deficiency, allergy,
CC asthma or cancer (including metastases). (I) are also used for
CC recombinant production of (II). (I) make it possible to tailor an immune
CC response to particular requirements, e.g. to direct a Th1-type helper
CC response to increase humoral or cellular responses (functioning as
CC adjuvant) to control B or T cell proliferation to induce immunogolobulin
CC synthesis or isotype switching. This sequence represents the HBsAg
CC (PreS2 plus S region) truncated peptide which is used in the method of
CC the invention.
XX
SQ Sequence 102 AA;

Query Match 100.0%; Score 61; DB 20; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 41 PLSSIFSRIGDP 52

RESULT 7
AAY93740
ID AAY93740 standard; Protein; 164 AA.
XX
AC AAY93740;
XX
DT 03-OCT-2000 (first entry)
XX
DE Amino acid sequence for pre-S peptide.
XX
KW Plant expression vector; antigen; hepatitis B surface antigen; HBsAg;
KW vaccine; ss.
XX
OS Synthetic.
XX
PN WO200037610-A2.
XX
PD 29-JUN-2000.
XX
PF 23-DEC-1999; 99WO-US31020.
XX
PR 23-DEC-1998; 98US-0113827.
XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (HEAL-) HEALTH RES INST.
XX
PI Mason HS, Thanavala Y, Arntzen CJ, Richter E;
XX
DR WPI; 2000-452181/39.
DR N-PSDB; AAA47000.
XX
PT New expression vector for transforming plants comprising two expression
PT cassettes useful for producing plant material comprising anti-hepatitis
PT B antibodies -
XX
PS Disclosure; Fig 5; 144pp; English.
XX
CC The specification describes a plant expression vector which comprises
CC two expression cassettes, the first comprising a polynucleotide encoding
CC an antigen and the second comprising a non-identical polynucleotide
CC encoding the same antigen. The antigen is especially a hepatitis B
CC surface antigen (HBsAg). The expression vector is used to transform
CC bacterial and plant cells to elicit the production of anti-hepatitis B
CC antibodies and are useful as vaccines. The present sequence represents
CC a pre-S peptide.
XX
SQ Sequence 164 AA;

Query Match 100.0%; Score 61; DB 21; Length 164;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 149 PLSSIFSRIGDP 160

RESULT 8
AAY90760
ID AAY90760 standard; Protein; 164 AA.
XX
AC AAY90760;
XX
DT 18-AUG-2000 (first entry)
XX
DE Hepatitis B virus preS protein sequence.
XX
KW Nuclear localisation signal; NLS; SV40; large T cell antigen;
KW fusion protein; diagnosis; detection; variant; vaccine; chaperone;
KW coprecipitate.
XX
OS Hepatitis B virus.
XX
PN WO200020606-A1.
XX
PD 13-APR-2000.

XX
PF 02-OCT-1998; 98WO-EP06298.
XX
PR 02-OCT-1998; 98WO-EP06298.
XX
PA (REIM/) REIMANN H.
PA (SCHI/) SCHIRMBECK R.
XX
PI Reimann H, Schirmbeck R;
XX
DR WPI; 2000-328936/28.
XX
PT Novel polynucleotide encoding a fusion protein which is stable in a
PT cell, useful for the production of peptides which coprecipitate
PT chaperones, by using truncated variants of the SV40 large T antigen
PT with an intact N-terminus -
XX
PS Example 1; Page 26; 57pp; English.
XX
CC The present invention describes a polynucleotide (I) encoding a fusion
CC protein (FP) (II) which is stable in a cell, and comprises a first
CC polypeptide (P1) and a second polypeptide (P2) which co-precipitates a
CC chaperone. The invention also provides methods for the production of
CC fusion proteins which comprise epitopes. Compositions comprising the
CC fusion proteins or epitopes of the invention are used as vaccines, and
CC for the production of antibodies. The methods provide an inexpensive
CC and efficient means of reliably expressing (poly)peptides comprising
CC epitopes. The present sequence represents the Hepatitis B virus (HBV)
CC preS protein sequence, which is used in an example from the present
CC invention.
XX
SQ Sequence 164 AA;

Query Match 100.0%; Score 61; DB 21; Length 164;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 150 PLSSIFSRIGDP 161

RESULT 9
AAR15617
ID AAR15617 standard; protein; 174 AA.
XX
AC AAR15617;
XX
DT 25-MAR-1992 (first entry)
XX
DE HBsAg pre-S region subtype ayw.
XX
KW T-cell epitope; vaccine; hepatitis B virus; antigen.
XX
OS Synthetic.
XX
PN WO9117768-A.
XX
PD 28-NOV-1991.
XX
PF 10-MAY-1991; 91WO-US03268.
XX
PR 11-MAY-1990; 90US-0522663.
XX
PA (SCRI) SCRIPPS CLINIC & RE.
XX
PI Milich DR, Thornton GB;
XX
DR WPI; 1991-369007/50.
XX
PT Hepatitis B virus surface antigen epitope(s) - useful as vaccines,
PT immunogens or diagnostic reagents
XX
PS Claim 1; Fig 1; 91pp; English.
XX
CC The amino acid sequence is that of a pre-S T cell epitope polypeptide
CC of the pre-S (2) region of hepatitis B surface antigen y (HBsAg/y).
CC It can be used to prime or vaccinate a host to induce responsiveness
CC to HBV vaccine. The T cell epitope polypeptides can also be used as
CC immunogens that prime T cells that respond to native HBsAg B cell
CC epitope polypeptide. The T cell epitope polypeptides are also useful
CC as substitutes for carrier immunogens such as KLH and are safe,
CC defined and T cell-active. In addition to their use as vaccines, the
CC polypeptides can be used as immunogens for prodn. of antibodies. See
CC also AAR15618-R15622.
XX
SQ Sequence 174 AA;

Query Match 100.0%; Score 61; DB 12; Length 174;
Best Local Similarity 100.0%; Pred. No. 0.00068;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 160 PLSSIFSRIGDP 171

RESULT 10
AAR23867
ID AAR23867 standard; protein; 174 AA.
XX
AC AAR23867;
XX
DT 11-NOV-1992 (first entry)
XX
DE Pre-S gene region translation product (1).
XX
KW env gene; hepatitis B virus; HBV; S-protein; hydrophilic;
KW hepadnavirus; immune response; antibody; envelope.
XX
OS Hepatitits B virus, subtype ayw.
XX
PN EP485361-A.
XX
PD 13-MAY-1992.
XX
PF 25-APR-1987; 87EP-0100663.
XX
PR 28-APR-1986; 86US-0856522.
XX
PA (CALY) CALIFORNIA INST TECHNOLOGY.
PA (NYBL-) NY BLOOD CENT INC.
XX
PI Kent SBH, Neurath AR;
XX
DR WPI; 1992-161100/20.
XX
PT Hepatitis B pre-S peptide immunogen and vaccine - for treatment
PT and diagnosis of hepatitis B
XX
PS Disclosure; Fig 2; 59pp; English.
XX
CC The sequences given in AAR23867 - AAR23871 are amino acid sequences
CC deduced from sequences of the pre-S portion of the env genes
CC corresponding to several hepatitis B virus (HBV) subtypes. These
CC proteins have properties distinct from those of the S-protein in
CC that they have high hydrophilicity and a high percentage of charged
CC residues, an abscence of Cys residues, the highest subtype-dependant
CC variability among HBV DNA gene products and little homology with
CC analogous sequences corresponding to nonhuman hepadnavirus. These
CC properties suggest that the pre-S gene coded portion of the HBV
CC envelope is expose to the surface of the virion, is a target for
CC the hosts immune response and is responsible for the host range of HBV
CC (limited to humans and some primates). Synthetic peptides and
CC antibodies against them, having predetermined specificity offer the
CC opportunity to explore the biological role of the pre-S protein moiety

CC of the HBV envelope. Portions of these amino acid sequences can be
CC used in a vaccine or in diagnostics for the detection of antigens and
CC antibodies, esp. those for the pre-S gene in sera of HBV infected
CC humans and certain animals, eg. chimpanzees.
XX
SQ Sequence 174 AA;

Query Match 100.0%; Score 61; DB 13; Length 174;
Best Local Similarity 100.0%; Pred. No. 0.00068;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 160 PLSSIFSRIGDP 171

RESULT 11
AAB10597
ID AAB10597 standard; Protein; 215 AA.
XX
AC AAB10597;
XX
DT 08-JAN-2001 (first entry)
XX
DE HBV fusion protein comprising HBcAg and RGD.
XX
KW Fusion protein; protein coat; virus-specific packaging signal; psi;
KW virus protein; cell permeability; cell-specific binding site; LHB;
KW large surface protein; core antigen; gene therapy.
XX
OS Hepatitis b virus.
OS Synthetic.
XX
PN WO200046376-A2.
XX
PD 10-AUG-2000.
XX
PF 04-FEB-2000; 2000WO-DE00363.
XX
PR 05-FEB-1999; 99DE-1004800.
XX
PA (HILD/) HILDT E.
XX
PI Hildt E, Hofschneider P;
XX
DR WPI; 2000-514959/46.
DR N-PSDB; AAA71735.
XX
PT Particle for cell-specific gene delivery, useful in gene therapy,
PT comprises nucleic acid in protein coat that includes a fusion protein
PT of viral protein, permeability peptide and cell-binding site -
XX
PS Claim 14; Fig 2; 34pp; German.
XX
CC This invention describes a novel particle (A), comprising a protein coat
CC with a fusion protein (FP), and, inside the coat, a nucleic acid (I)
CC including the sequence for a virus-specific packaging signal (psi) and a
CC structural gene. FP contains a virus protein (VP), a peptide (P) that
CC mediates cell permeability and a heterologous cell-specific binding site
CC (RGD). The invention also describes (1) producing (A) in which FP
CC contains an LHBs (large surface protein of hepatitis B virus (HBV)) and
CC RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),
CC (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector
CC containing the DNA of (d). The products of the invention are used in gene
CC therapy of cells and tissues, in vivo or ex vivo. This sequence
CC represents a fusion protein which is described in the method of the
CC invention.
XX
SQ Sequence 215 AA;

Query Match 100.0%; Score 61; DB 21; Length 215;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 2 PLSSIFSRIGDP 13

RESULT 12
AAR27471
ID AAR27471 standard; Protein; 281 AA.
XX
AC AAR27471;
XX
DT 24-FEB-1993 (first entry)
XX
DE spsAg protein.
XX
KW Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
KW vaccinia virus; H6; early/late; promoter; NYVAC; recombinant; HBV L;
KW large pre-S antigen; lpsAg; fusion protein; pre-S region; S12/core;
KW S1; S2; Copenhagen vaccine strain; vaccinia virus; virulence factor;
KW deletion loci; recipient loci.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..55
FT /label= S2
FT Region 56..281
FT /label= S
XX
PN WO9215672-A.
XX
PD 17-SEP-1992.
XX
PF 09-MAR-1992; 92WO-US01906.
XX
PR 07-MAR-1991; 91US-0666056.
PR 11-JUN-1991; 91US-0713967.
PR 06-MAR-1992; 92US-0847951.
XX
PA (VIRO-) VIROGENETICS CORP.
XX
PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP;
PI Limbach KJ, Norton EK, Paoletti E, Perkus ME, Pincus SE;
PI Riviere M, Tartaglia J, Taylor J;
XX
DR WPI; 1992-331718/40.
DR N-PSDB; AAQ29103.
XX
PT Vaccine comprises recombinant, attenuated pox-virus - use for
PT vaccinating against viral infections such as rabies, hepatitis B,
PT HIV, HSV, EBV, CMV, mumps etc.
XX
PS Disclosure; Fig 9; 456pp; English.
XX
CC The sequence given is encoded by an expression cassette which
CC consists of the hepatitis B virus (HBV) M protein (small pre-S
CC antigen, spsAg) gene precisely linked to a modified synthetic
CC vaccinia virus H6 early/late promoter. This DNA sequence was used in
CC the construction of a NYVAC recombinant expressing the HBV gene.
CC Other HBV genes were also used in the construction. These were HBV
CC L (large pre-S antigen, lpsAg) and a fusion protein composed of the
CC entire pre-S region (S12/core, S1 + S2). Each of these gene sequences
CC were inserted individually into three different sites of NYVAC
CC separated by from each other by large regions of vaccinia DNA
CC containing essential genes. NYVAC is a Copenhagen vaccine strain of
CC vaccinia virus which has been modified by deletion of six non-essential
CC regions of the genome encoding known or potential virulence factors.
CC The deletion loci were engineered as recipient loci for the insertion
CC of foreign genes. The spacing of the three inserted sequences ensured
CC that any recombination that did occur would lead to disruption of the
CC vaccinia genome and would cause unviable vaccinia virus. See also
CC AAQ35501-864.

XX
SQ Sequence 281 AA;

Query Match 100.0%; Score 61; DB 13; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 41 PLSSIFSRIGDP 52

RESULT 13
AAY32835
ID AAY32835 standard; Protein; 281 AA.
XX
AC AAY32835;
XX
DT 29-OCT-1999 (first entry)
XX
DE HBsAg PreS2-S region protein sequence.
XX
KW HBsAg; PreS2-S; recombinant antigen library; disease-related antigen;
KW multivalent antigenic polypeptide production; infection; allergen;
KW asthma; autoimmune disease; rheumatoid arthritis; diabetes; therapy;
KW multiple sclerosis; inflammatory condition; cancer; contraception;
KW immune response; hepatitis b surface antigen.
XX
OS Hepatitis b virus.
XX
PN WO9941383-A1.
XX
PD 19-AUG-1999.
XX
PF 10-FEB-1999; 99WO-US02944.
XX
PR 23-OCT-1998; 98US-0105509.
PR 11-FEB-1998; 98US-0021769.
PR 11-FEB-1998; 98US-0074294.
XX
PA (MAXY-) MAXYGEN INC.
XX
PI Bass SH, Howard R, Punnonen J, Stemmer WPC, Whalen RG;
XX
DR WPI; 1999-518452/43.
DR N-PSDB; AAZ10968.
XX
PT Recombinant multivalent antigenic polypeptide produced by
PT recombining nucleic acid sequences and screening, used in vaccines
PT against e.g. infections and cancer
XX
PS Example 14; Fig 17; 153pp; English.
XX
CC This sequence is the hepatitis B virus (HBV) surface antigen (HBsAg)
CC PreS2-S region. This sequence was used to create a recombinant antigen
CC library. The library comprises recombinant nucleic acids encoding
CC antigenic polypeptides and is produced by recombination of at least two
CC forms of nucleic acid, differing by at least two nucleotides, encoding a
CC disease-related antigenic polypeptide. The library can be used to produce
CC a recombinant multivalent antigenic polypeptides of the invention, that
CC contains at least two antigenic determinants (AD) from different
CC polypeptides. The multivalent antigenic polypeptides are used in vaccines
CC to induce a protective or therapeutic response to a wide variety of
CC infectious agents (bacteria, viruses, parasites, including Plasmodium
CC falciparum); allergens; asthma; autoimmune disease (e.g. rheumatoid
CC arthritis, diabetes, multiple sclerosis); other inflammatory conditions
CC and cancer, also, where directed against sperm antigens,they can be used
CC for contraception. The multivalent peptides can be evolved to induce an
CC optimised immune response against a wide variety of antigens,
CC particularly a broad spectrum response to many different strains of a
CC pathogen, including strains that are likely to appear in the future.
XX
SQ Sequence 281 AA;

Query Match 100.0%; Score 61; DB 20; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 41 PLSSIFSRIGDP 52

RESULT 14
AAR27472
ID AAR27472 standard; Protein; 389 AA.
XX
AC AAR27472;
XX
DT 24-FEB-1993 (first entry)
XX
DE lpsAg protein.
XX
KW Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
KW vaccinia virus; hemorraghic region; u; promoter; NYVAC; recombinant;
KW HBV L; large pre-S antigen; lpsAg; fusion protein; pre-S region;
KW S12/core; S1; S2; Copenhagen vaccine strain; vaccinia virus;
KW virulence factor; deletion loci; recipient loci.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT region 1..108
FT /label= S1
FT region 109..163
FT /label= S2
FT misc_RNA 164..389
FT /label= S
XX
PN WO9215672-A.
XX
PD 17-SEP-1992.
XX
PF 09-MAR-1992; 92WO-US01906.
XX
PR 07-MAR-1991; 91US-0666056.
PR 11-JUN-1991; 91US-0713967.
PR 06-MAR-1992; 92US-0847951.
XX
PA (VIRO-) VIROGENETICS CORP.
XX
PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP;
PI Limbach KJ, Norton EK, Paoletti E, Perkus ME, Pincus SE;
PI Riviere M, Tartaglia J, Taylor J;
XX
DR WPI; 1992-331718/40.
DR N-PSDB; AAQ29104.
XX
PT Vaccine comprises recombinant, attenuated pox-virus - use for
PT vaccinating against viral infections such as rabies, hepatitis B,
PT HIV, HSV, EBV, CMV, mumps etc.
XX
PS Disclosure; Fig 11; 456pp; English.
XX
CC The sequence given is encoded by an expression vector which comprises
CC the hepatitis B virus (HBV) L protein (large pre-S antigen, lpsAg)
CC gene linked to the cowpox hemorraghic region (u) promoter. This DNA
CC sequence was used in the construction of a NYVAC recombinant
CC expressing the HBV gene. Other HBV genes were also used in the
CC construction. These were HBV M protein (small pre-S antigen, spsAg)
CC and a fusion protein composed of the entire pre-S region (S12/core,
CC S1 + S2). Each of these gene sequences were inserted individually
CC into three different sites of NYVAC separated by from each other by
CC large regions of vaccinia DNA containing essential genes. NYVAC is a
CC Copenhagen vaccine strain of vaccinia virus which has been modified by
CC deletion of six non-essential regions of the genome encoding known or

```
CC   potential virulence factors.  The deletion loci were engineered as
CC   recipient loci for the insertion of foreign genes.  The spacing of the
CC   three inserted sequences ensured that any recombination that did occur
CC   would lead to disruption of the vaccinia genome and would cause
CC   unviable vaccinia virus.  See also AAQ35501-864.
XX
SQ   Sequence   389 AA;

  Query Match             100.0%;  Score 61;  DB 13;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.0016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 15
AAR27474
ID   AAR27474 standard; Protein; 389 AA.
XX
AC   AAR27474;
XX
DT   24-FEB-1993  (first entry)
XX
DE   I3L promoter/S12/core gene.
XX
KW   Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
KW   vaccinia virus; Amsacta moorei entomopoxvirus; AmEPV; 42 kD; promoter;
KW   NYVAC; recombinant; HBV L; large pre-S antigen; lpsAg; fusion protein;
KW   pre-S region; S12/core; S1; S2; Copenhagen vaccine strain;
KW   vaccinia virus; virulence factor; deletion loci; recipient loci.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1..108
FT                   /label= S1
FT   Region          109..163
FT                   /label= S2
FT   Region          164..389
FT                   /label= S
XX
PN   WO9215672-A.
XX
PD   17-SEP-1992.
XX
PF   09-MAR-1992;   92WO-US01906.
XX
PR   07-MAR-1991;   91US-0666056.
PR   11-JUN-1991;   91US-0713967.
PR   06-MAR-1992;   92US-0847951.
XX
PA   (VIRO-) VIROGENETICS CORP.
XX
PI   Cox WI,  De Taisne C,  Francis J,  Gettig RR,  Johnson GP;
PI   Limbach KJ,  Norton EK,  Paoletti E,  Perkus ME,  Pincus SE;
PI   Riviere M,  Tartaglia J,  Taylor J;
XX
DR   WPI; 1992-331718/40.
DR   N-PSDB; AAQ29106.
XX
PT   Vaccine comprises recombinant, attenuated pox-virus - use for
PT   vaccinating against viral infections such as rabies, hepatitis B,
PT   HIV, HSV, EBV, CMV, mumps etc.
XX
PS   Disclosure; Fig 15; 456pp; English.
XX
CC   The sequence given is encoded by an expression cassette which
CC   comprises the hepatitis B virus (HBV) L protein (large pre-S antigen,
CC   lpsAg) gene which is precisely linked to the Amsacta moorei
CC   entomopoxvirus (AmEPV) 42 kD promoter.  This DNA sequence was used in
CC   the construction of a NYVAC recombinant expressing the HBV gene.
CC   Other HBV genes were also used in the construction.  These were HBV M
CC   protein (small pre-S antigen, spsAg) and a fusion protein composed of
CC   the entire pre-S region (S12/core, S1 + S2).  Each of these gene
CC   sequences were inserted individually into three different sites of
CC   NYVAC separated by from each other by large regions of vaccinia DNA
CC   containing essential genes.  NYVAC is a Copenhagen vaccine strain of
CC   vaccinia virus which has been modified by deletion of six non-essential
CC   regions of the genome encoding known or potential virulence factors.
CC   The deletion loci were engineered as recipient loci for the insertion
CC   of foreign genes.  The spacing of the three inserted sequences ensured
CC   that any recombination that did occur would lead to disruption of the
CC   vaccinia genome and would cause unviable vaccinia virus.  See also
CC   AAQ35501-864.
XX
SQ   Sequence   389 AA;

  Query Match             100.0%;  Score 61;  DB 13;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.0016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


Search completed: February  3, 2003, 09:23:52
Job time : 2.92335 secs
```

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:23:57 ; Search time 8.57143 Seconds
(without alignments)
902.626 Million cell updates/sec

Title: US-09-890-752A-20
Perfect score: 61
Sequence: 1 PLSSIFSRIGDP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61	100.0	12	22	US-09-830-981-2	Sequence 2, Appli
2	61	100.0	12	22	US-09-830-981-4	Sequence 4, Appli
3	61	100.0	12	22	US-09-830-981-12	Sequence 12, Appl
4	61	100.0	12	22	US-09-830-981-14	Sequence 14, Appl
5	61	100.0	12	22	US-09-830-981-16	Sequence 16, Appl
6	61	100.0	12	22	US-09-830-981-18	Sequence 18, Appl
7	61	100.0	12	22	US-09-830-981-20	Sequence 20, Appl
8	61	100.0	12	22	US-09-890-752A-20	Sequence 20, Appl
9	61	100.0	12	24	US-10-077-555-9	Sequence 9, Appli
10	61	100.0	12	25	US-10-144-549-12	Sequence 12, Appl
11	61	100.0	12	26	US-10-226-956-284	Sequence 284, App
12	61	100.0	102	14	US-09-021-769-45	Sequence 45, Appl
13	61	100.0	102	16	US-09-248-716-6	Sequence 6, Appli
14	61	100.0	102	21	US-09-724-890-45	Sequence 45, Appl
15	61	100.0	164	1	PCT-US99-31020-41	Sequence 41, Appl
16	61	100.0	164	18	US-09-471-573A-41	Sequence 41, Appl
17	61	100.0	215	22	US-09-890-752A-2	Sequence 2, Appli
18	61	100.0	281	3	US-07-847-951A-214	Sequence 214, App
19	61	100.0	281	4	US-08-036-218-212	Sequence 212, App
20	61	100.0	281	14	US-09-021-769-6	Sequence 6, Appli
21	61	100.0	281	16	US-09-247-890-12	Sequence 12, Appl
22	61	100.0	281	21	US-09-724-648-12	Sequence 12, Appl
23	61	100.0	281	21	US-09-724-890-6	Sequence 6, Appli
24	61	100.0	281	21	US-09-724-969-12	Sequence 12, Appl
25	61	100.0	389	3	US-07-847-951A-216	Sequence 216, App
26	61	100.0	389	3	US-07-847-951A-219	Sequence 219, App
27	61	100.0	389	4	US-08-036-218-214	Sequence 214, App
28	61	100.0	389	4	US-08-036-218-217	Sequence 217, App
29	61	100.0	389	7	US-08-360-107-116	Sequence 116, App
30	61	100.0	389	8	US-08-470-896-106	Sequence 106, App
31	61	100.0	389	8	US-08-471-913-106	Sequence 106, App
32	61	100.0	389	8	US-08-475-668-106	Sequence 106, App
33	61	100.0	389	8	US-08-484-223-106	Sequence 106, App
34	61	100.0	389	8	US-08-484-223A-106	Sequence 106, App
35	61	100.0	389	8	US-08-485-546-106	Sequence 106, App
36	61	100.0	389	8	US-08-485-546A-106	Sequence 106, App
37	61	100.0	389	8	US-08-485-551-106	Sequence 106, App
38	61	100.0	389	8	US-08-487-266-106	Sequence 106, App
39	61	100.0	389	8	US-08-487-266A-106	Sequence 106, App
40	61	100.0	389	8	US-08-487-355-106	Sequence 106, App
41	61	100.0	389	8	US-08-487-355A-106	Sequence 106, App
42	61	100.0	389	13	US-08-919-600-106	Sequence 106, App
43	61	100.0	389	19	US-09-502-445-106	Sequence 106, App
44	61	100.0	389	26	US-10-267-682-106	Sequence 106, App
45	61	100.0	389	26	US-10-267-748-106	Sequence 106, App

ALIGNMENTS

```
RESULT 1
US-09-830-981-2
; Sequence 2, Application US/09830981
; GENERAL INFORMATION:
;  APPLICANT: Eberhard Hildt
;  APPLICANT:  Stephanie Schmidt
;  TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
;  FILE REFERENCE: 033392-001
;  CURRENT APPLICATION NUMBER: US/09/830,981
;  CURRENT FILING DATE:  2002-04-15
;  PRIOR APPLICATION NUMBER: PCT/DE99/03506
;  PRIOR FILING DATE: 1999-11-03
;  PRIOR APPLICATION NUMBER: DE 198 50 718.6
;  PRIOR FILING DATE: 1998-11-03
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Cell permeability mediating polypeptide
US-09-830-981-2

  Query Match           100.0%;  Score 61;  DB 22;  Length 12;
  Best Local Similarity  100.0%;  Pred. No. 0.00042;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db      1 PLSSIFSRIGDP 12


RESULT 2
US-09-830-981-4
; Sequence 4, Application US/09830981
; GENERAL INFORMATION:
;  APPLICANT: Eberhard Hildt
;  APPLICANT:  Stephanie Schmidt
;  TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
;  FILE REFERENCE: 033392-001
;  CURRENT APPLICATION NUMBER: US/09/830,981
;  CURRENT FILING DATE:  2002-04-15
;  PRIOR APPLICATION NUMBER: PCT/DE99/03506
;  PRIOR FILING DATE: 1999-11-03
;  PRIOR APPLICATION NUMBER: DE 198 50 718.6
;  PRIOR FILING DATE: 1998-11-03
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: Hepadnavirus
US-09-830-981-4

  Query Match             100.0%;  Score 61;  DB 22;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.00042;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db      1 PLSSIFSRIGDP 12


RESULT 3
US-09-830-981-12
; Sequence 12, Application US/09830981
; GENERAL INFORMATION:
;  APPLICANT: Eberhard Hildt
;  APPLICANT:  Stephanie Schmidt
;  TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
;  FILE REFERENCE: 033392-001
;  CURRENT APPLICATION NUMBER: US/09/830,981
;  CURRENT FILING DATE:  2002-04-15
;  PRIOR APPLICATION NUMBER: PCT/DE99/03506
;  PRIOR FILING DATE: 1999-11-03
;  PRIOR APPLICATION NUMBER: DE 198 50 718.6
;  PRIOR FILING DATE: 1998-11-03
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: Hepadnavirus
US-09-830-981-12

  Query Match             100.0%;  Score 61;  DB 22;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.00042;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db      1 PLSSIFSRIGDP 12


RESULT 4
US-09-830-981-14
; Sequence 14, Application US/09830981
; GENERAL INFORMATION:
;  APPLICANT: Eberhard Hildt
;  APPLICANT:  Stephanie Schmidt
;  TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
;  FILE REFERENCE: 033392-001
;  CURRENT APPLICATION NUMBER: US/09/830,981
;  CURRENT FILING DATE:  2002-04-15
;  PRIOR APPLICATION NUMBER: PCT/DE99/03506
;  PRIOR FILING DATE: 1999-11-03
;  PRIOR APPLICATION NUMBER: DE 198 50 718.6
;  PRIOR FILING DATE: 1998-11-03
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: Hepadnavirus
US-09-830-981-14

  Query Match             100.0%;  Score 61;  DB 22;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.00042;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db      1 PLSSIFSRIGDP 12


RESULT 5
US-09-830-981-16
; Sequence 16, Application US/09830981
; GENERAL INFORMATION:
;  APPLICANT: Eberhard Hildt
;  APPLICANT:  Stephanie Schmidt
;  TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
;  FILE REFERENCE: 033392-001
;  CURRENT APPLICATION NUMBER: US/09/830,981
;  CURRENT FILING DATE:  2002-04-15
;  PRIOR APPLICATION NUMBER: PCT/DE99/03506
;  PRIOR FILING DATE: 1999-11-03
;  PRIOR APPLICATION NUMBER: DE 198 50 718.6
;  PRIOR FILING DATE: 1998-11-03
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: Hepadnavirus
US-09-830-981-16

  Query Match             100.0%;  Score 61;  DB 22;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.00042;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db      1 PLSSIFSRIGDP 12


RESULT 6
US-09-830-981-18
; Sequence 18, Application US/09830981
; GENERAL INFORMATION:
;  APPLICANT: Eberhard Hildt
;  APPLICANT:  Stephanie Schmidt
;  TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
;  FILE REFERENCE: 033392-001
;  CURRENT APPLICATION NUMBER: US/09/830,981
;  CURRENT FILING DATE:  2002-04-15
;  PRIOR APPLICATION NUMBER: PCT/DE99/03506
;  PRIOR FILING DATE: 1999-11-03
;  PRIOR APPLICATION NUMBER: DE 198 50 718.6
;  PRIOR FILING DATE: 1998-11-03
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
```

; LENGTH: 12
; TYPE: PRT
; ORGANISM: Hepadnavirus
US-09-830-981-18

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 1 PLSSIFSRIGDP 12

RESULT 7
US-09-830-981-20
; Sequence 20, Application US/09830981
; GENERAL INFORMATION:
; APPLICANT: Eberhard Hildt
; APPLICANT: Stephanie Schmidt
; TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
; FILE REFERENCE: 033392-001
; CURRENT APPLICATION NUMBER: US/09/830,981
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: PCT/DE99/03506
; PRIOR FILING DATE: 1999-11-03
; PRIOR APPLICATION NUMBER: DE 198 50 718.6
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Hepadnavirus
US-09-830-981-20

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 1 PLSSIFSRIGDP 12

RESULT 8
US-09-890-752A-20
; Sequence 20, Application US/09890752A
; GENERAL INFORMATION:
; APPLICANT: Hildt, Eberhard
; APPLICANT: Hofschneider, Peter
; TITLE OF INVENTION: Particles for Gene Therapy
; FILE REFERENCE: 107070-120 (VOS-013)
; CURRENT APPLICATION NUMBER: US/09/890,752A
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: PCT/DE00/00363
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: DE 199 04 800.2
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: cell permeability-mediating peptide
US-09-890-752A-20

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 1 PLSSIFSRIGDP 12

RESULT 9
US-10-077-555-9
; Sequence 9, Application US/10077555
; GENERAL INFORMATION:
; APPLICANT: Wang, Rong-fu
; TITLE OF INVENTION: Use of Cell-Penetrating Peptides to Generate Antitumor Immunity
; FILE REFERENCE: P02373US1/10200806
; CURRENT APPLICATION NUMBER: US/10/077,555
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268,687
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-077-555-9

Query Match 100.0%; Score 61; DB 24; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 1 PLSSIFSRIGDP 12

RESULT 10
US-10-144-549-12
; Sequence 12, Application US/10144549
; GENERAL INFORMATION:
; APPLICANT: GeneShuttle Biopharm, Inc.
; APPLICANT: Hwu , Paul L.
; TITLE OF INVENTION: A NEW FUSION PROTEIN FOR USE AS VECTOR
; FILE REFERENCE: MBHB 02-340
; CURRENT APPLICATION NUMBER: US/10/144,549
; CURRENT FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Hepatitis B virus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; OTHER INFORMATION: The HBV PreS2 antigen consisting of the translocation motif from
; OTHER INFORMATION: residues 41-52.
US-10-144-549-12

Query Match 100.0%; Score 61; DB 25; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 1 PLSSIFSRIGDP 12

RESULT 11
US-10-226-956-284
; Sequence 284, Application US/10226956
; GENERAL INFORMATION:
; APPLICANT: Brophy, Colleen
; APPLICANT: Komalavilas, Padmini
; APPLICANT: Panitch, Alyssa
; APPLICANT: Joshi, Lokesh
; APPLICANT: Seal, Brandon L.

```
;  TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
;  FILE REFERENCE: ASU-1061-US
;  CURRENT APPLICATION NUMBER: US/10/226,956
;  CURRENT FILING DATE:  2002-08-23
;  PRIOR APPLICATION NUMBER: 60/314,535
;  PRIOR FILING DATE: 2001-08-23
;  NUMBER OF SEQ ID NOS: 320
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 284
;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic peptide
US-10-226-956-284

  Query Match             100.0%;  Score 61;  DB 26;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.00042;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db       1 PLSSIFSRIGDP 12


RESULT 12
US-09-021-769-45
; Sequence 45, Application US/09021769
;  GENERAL INFORMATION:
;    APPLICANT:  Punnonen, Juha
;    APPLICANT:  Stemmer, Willem P.C.
;    APPLICANT:  Whalen, Robert Gerald
;    APPLICANT:  Howard, Russell
;    TITLE OF INVENTION:  Methods and Compositions for Vaccine
;    TITLE OF INVENTION:  Engineering
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/021,769
;      FILING DATE:  11-FEB-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, Timothy L.
;      REGISTRATION NUMBER:  35,367
;      REFERENCE/DOCKET NUMBER:  018097-021500US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  102 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..102
;      OTHER INFORMATION:  /notes "truncated Hepatitis B surface
;      OTHER INFORMATION:  antigen (HBsAg) (PreS2 plus S regions)
;      OTHER INFORMATION:  polypeptide"
US-09-021-769-45

  Query Match             100.0%;  Score 61;  DB 14;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 0.0044;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db      41 PLSSIFSRIGDP 52


RESULT 13
US-09-248-716-6
; Sequence 6, Application US/09248716
; GENERAL INFORMATION:
;  APPLICANT:  Punnonen, Juha
;  APPLICANT:  Stemmer, Willem P.C.
;  APPLICANT:  Whalen, Robert Gerald
;  APPLICANT:  Howard, Russell
;  APPLICANT:  Maxygen, Inc.
;  TITLE OF INVENTION:  Optimization of Immunomodulatory Properties of Genetic
;  TITLE OF INVENTION:  Vaccines
;  FILE REFERENCE:  018097-030300US
;  CURRENT APPLICATION NUMBER: US/09/248,716
;  CURRENT FILING DATE:  1999-02-10
;  PRIOR APPLICATION NUMBER:  US 60/074,294
;  PRIOR FILING DATE:  1998-02-11
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 102
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:truncated
;   OTHER INFORMATION: hepatitis B surface antigen (HBsAg) (PreS2 plus S
;   OTHER INFORMATION: regions)
US-09-248-716-6

  Query Match             100.0%;  Score 61;  DB 16;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 0.0044;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db      41 PLSSIFSRIGDP 52


RESULT 14
US-09-724-890-45
; Sequence 45, Application US/09724890
;   GENERAL INFORMATION:
;        APPLICANT: Punnonen, Juha
;                   Stemmer, Willem P.C.
;                   Whalen, Robert Gerald
;                   Howard, Russell
;        TITLE OF INVENTION: Methods and Compositions for Vaccine
;                            Engineering
;        NUMBER OF SEQUENCES: 51
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/724,890
;             FILING DATE: 28-Nov-2000
```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,769
; FILING DATE: 11-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 018097-021500US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..102
; OTHER INFORMATION: /note= "truncated Hepatitis B
; surface
; antigen (HBsAg) (PreS2 plus S
; regions)
; polypeptide"
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-724-890-45

Query Match 100.0%; Score 61; DB 21; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 41 PLSSIFSRIGDP 52

RESULT 15
PCT-US99-31020-41
; Sequence 41, Application PC/TUS9931020
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Inst. for Plant Res. at Cornell
; TITLE OF INVENTION: Expression of Immunogenic Hepatitis B Surface Antigens
; TITLE OF INVENTION: in Transgenic Plants
; FILE REFERENCE: 4868/85428
; CURRENT APPLICATION NUMBER: PCT/US99/31020
; CURRENT FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Plant-optimized
; OTHER INFORMATION: pre S1/S2.
PCT-US99-31020-41

Query Match 100.0%; Score 61; DB 1; Length 164;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSSIFSRIGDP 12
||||||||||||
Db 149 PLSSIFSRIGDP 160

Search completed: February 3, 2003, 09:34:08
Job time : 8.57143 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:12 ; Search time 1.60976 Seconds
(without alignments)
1535.988 Million cell updates/sec

Title: US-09-890-752A-20
Perfect score: 61
Sequence: 1 PLSSIFSRIGDP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61	100.0	57	12	Q91IC1	Q91ic1 hepatitis b
2	61	100.0	163	12	Q9WPI9	Q9wpi9 hepatitis b
3	61	100.0	281	12	O41750	O41750 hepatitis b
4	61	100.0	281	12	Q8QXP8	Q8qxp8 hepatitis b
5	61	100.0	378	12	Q67944	Q67944 hepatitis b
6	61	100.0	378	12	Q67954	Q67954 hepatitis b
7	61	100.0	383	12	Q96837	Q96837 hepatitis b
8	61	100.0	383	12	Q96840	Q96840 hepatitis b
9	61	100.0	389	12	O90772	O90772 hepatitis b
10	61	100.0	389	12	Q9WMX3	Q9wmx3 hepatitis b
11	61	100.0	389	12	Q91C55	Q91c55 hepatitis b
12	61	100.0	389	12	Q91C52	Q91c52 hepatitis b
13	61	100.0	389	12	Q67875	Q67875 hepatitis b
14	61	100.0	389	12	Q8QXP9	Q8qxp9 hepatitis b
15	59	96.7	49	12	Q91IC3	Q91ic3 hepatitis b
16	59	96.7	55	12	Q9IXF4	Q9ixf4 hepatitis b
17	59	96.7	55	12	Q9IXF2	Q9ixf2 hepatitis b
18	59	96.7	55	12	Q9IXF1	Q9ixf1 hepatitis b
19	59	96.7	55	12	Q9IXE9	Q9ixe9 hepatitis b
20	59	96.7	55	12	Q9IXE8	Q9ixe8 hepatitis b
21	59	96.7	55	12	Q9IXE6	Q9ixe6 hepatitis b
22	59	96.7	55	12	Q9IXE2	Q9ixe2 hepatitis b
23	59	96.7	55	12	Q9IXE0	Q9ixe0 hepatitis b
24	59	96.7	55	12	Q9IXD9	Q9ixd9 hepatitis b
25	59	96.7	55	12	Q9IXD7	Q9ixd7 hepatitis b
26	59	96.7	55	12	Q9IXD6	Q9ixd6 hepatitis b
27	59	96.7	55	12	Q9IXD4	Q9ixd4 hepatitis b
28	59	96.7	55	12	Q9IXD3	Q9ixd3 hepatitis b
29	59	96.7	55	12	Q9IXD1	Q9ixd1 hepatitis b
30	59	96.7	55	12	Q9IXD0	Q9ixd0 hepatitis b
31	59	96.7	55	12	Q9IXC8	Q9ixc8 hepatitis b
32	59	96.7	55	12	Q9IXC7	Q9ixc7 hepatitis b
33	59	96.7	55	12	Q9IXC5	Q9ixc5 hepatitis b
34	59	96.7	55	12	Q9IXC4	Q9ixc4 hepatitis b
35	59	96.7	55	12	Q9IXC2	Q9ixc2 hepatitis b
36	59	96.7	55	12	Q9IXB5	Q9ixb5 hepatitis b
37	59	96.7	55	12	Q9IXB3	Q9ixb3 hepatitis b
38	59	96.7	55	12	Q9IXB2	Q9ixb2 hepatitis b
39	59	96.7	55	12	Q9IXB0	Q9ixb0 hepatitis b
40	59	96.7	55	12	Q9IXA9	Q9ixa9 hepatitis b
41	59	96.7	55	12	Q9IXA7	Q9ixa7 hepatitis b
42	59	96.7	55	12	Q9IXA6	Q9ixa6 hepatitis b
43	59	96.7	55	12	Q9IXA4	Q9ixa4 hepatitis b
44	59	96.7	55	12	Q9IXA3	Q9ixa3 hepatitis b
45	59	96.7	55	12	Q9IXA1	Q9ixa1 hepatitis b

ALIGNMENTS

```
RESULT 1
Q91IC1
ID   Q91IC1      PRELIMINARY;      PRT;    57 AA.
AC   Q91IC1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Pre-S2 protein (Fragment).
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N314-1;
RA   Huangfu J., Dong J., Deng H.;
RT   "Preliminary Study on PreS2 Region of Hepatitis B Virus.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF393223; AAK84362.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
FT   NON_TER      57     57
SQ   SEQUENCE   57 AA;  6124 MW;  8D756AF65FF849CF CRC64;

  Query Match            100.0%;  Score 61;  DB 12;  Length 57;
  Best Local Similarity  100.0%;  Pred. No. 0.00019;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db     41 PLSSIFSRIGDP 52


RESULT 2
Q9WPI9
ID   Q9WPI9      PRELIMINARY;      PRT;   163 AA.
AC   Q9WPI9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
```

```
DE   Large S protein (Fragment).
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MF;
RA   Kidd-Ljunggren K.;
RT   "Nosocomial Transmission of Hepatitis B Virus Through Multiple Dose
RT   Vials.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF136461; AAD44711.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
FT   NON_TER     163    163
SQ   SEQUENCE   163 AA;  17331 MW;  38D20BE460F0BD24 CRC64;

  Query Match             100.0%;  Score 61;  DB 12;  Length 163;
  Best Local Similarity   100.0%;  Pred. No. 0.00061;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 3
O41750
ID   O41750      PRELIMINARY;      PRT;   281 AA.
AC   O41750;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Middle S protein.
GN   PRES2/S.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E08;
RX   MEDLINE=97368435; PubMed=9225049;
RA   Bowyer S., van Staden L., Kew M.C., Sim J.G.;
RT   "A unique segment of the hepatitis B virus group A genotype identified
RT   in isolates from South Africa.";
RL   J. Gen. Virol. 78:1719-1729(1997).
DR   EMBL; U87737; AAC58018.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   281 AA;  31153 MW;  85D90E7B5F2C9289 CRC64;

  Query Match             100.0%;  Score 61;  DB 12;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.0011;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db      41 PLSSIFSRIGDP 52


RESULT 4
Q8QXP8
ID   Q8QXP8      PRELIMINARY;      PRT;   281 AA.
AC   Q8QXP8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Middle surface antigen.
GN   PRES2.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kay A.C.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Jeantet D., Chemin I., Mandrand B., Zoulim F., Trepo C., Kay A.;
RT   "Characterization of two Hepatitis B virus populations isolated from a
RT   Hepatitis B surface antigen-negative patient.";
RL   Hepatology 0:0-0(0).
DR   EMBL; AJ344117; CAC87014.1; -.
SQ   SEQUENCE   281 AA;  31266 MW;  37234BBD949B3B03 CRC64;

  Query Match             100.0%;  Score 61;  DB 12;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.0011;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db      41 PLSSIFSRIGDP 52


RESULT 5
Q67944
ID   Q67944      PRELIMINARY;      PRT;   378 AA.
AC   Q67944;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Surface proteins.
GN   PRES1/PRES2.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X85254; CAA59510.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   378 AA;  41257 MW;  D4370B06FFB625D5 CRC64;

  Query Match             100.0%;  Score 61;  DB 12;  Length 378;
  Best Local Similarity   100.0%;  Pred. No. 0.0015;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 6
Q67954
ID   Q67954      PRELIMINARY;      PRT;   378 AA.
AC   Q67954;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   HBV surface proteins (Fragment).
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lai M.E., Mazzoleni A.P., Balestrieri A.;
RT   "Sequence analysis of HBV genomes isolated from patients with HBsAg
RT   chronic liver disease.";
RL   Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X77309; CAA54515.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
FT   NON_TER     378    378
SQ   SEQUENCE   378 AA;  41356 MW;  FC18839606E121CD CRC64;
```

```
  Query Match           100.0%;  Score 61;  DB 12;  Length 378;
  Best Local Similarity  100.0%;  Pred. No. 0.0015;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db    149 PLSSIFSRIGDP 160


RESULT 7
Q96837
ID   Q96837      PRELIMINARY;      PRT;   383 AA.
AC   Q96837;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Large S protein (Fragment).
GN   PRE-S/S.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRH;
RX   MEDLINE=97049068; PubMed=8893798;
RA   Moraes M.T., Gomes S.A., Niel C.;
RT   "Sequence analysis of pre-S/S gene of hepatitis B virus strains of
RT   genotypes A, D, and F isolated in Brazil.";
RL   Arch. Virol. 141:1767-1773(1996).
DR   EMBL; U55224; AAB47470.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
FT   NON_TER     383     383
SQ   SEQUENCE   383 AA;  41942 MW;  6E101F3B7471B6B1 CRC64;

  Query Match           100.0%;  Score 61;  DB 12;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.0015;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db    149 PLSSIFSRIGDP 160


RESULT 8
Q96840
ID   Q96840      PRELIMINARY;      PRT;   383 AA.
AC   Q96840;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Large S protein (Fragment).
GN   PRE-S/S.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRJ;
RX   MEDLINE=97049068; PubMed=8893798;
RA   Moraes M.T., Gomes S.A., Niel C.;
RT   "Sequence analysis of pre-S/S gene of hepatitis B virus strains of
RT   genotypes A, D, and F isolated in Brazil.";
RL   Arch. Virol. 141:1767-1773(1996).
DR   EMBL; U55227; AAB47473.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
FT   NON_TER     383     383
SQ   SEQUENCE   383 AA;  42018 MW;  341AB53B7911BBDC CRC64;

  Query Match           100.0%;  Score 61;  DB 12;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.0015;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db    149 PLSSIFSRIGDP 160


RESULT 9
O90772
ID   O90772      PRELIMINARY;      PRT;   389 AA.
AC   O90772;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Surface protein.
GN   S GENE.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=REF2;
RX   MEDLINE=98362187; PubMed=9696878;
RA   Grethe S., Monazahian M., Boehme I., Thomssen R.;
RT   "Characterization of unusual escape variants of hepatitis B virus
RT   isolated from a hepatitis B surface antigen-negative subject.";
RL   J. Virol. 72:7692-7696(1998).
DR   EMBL; AJ003116; CAA05872.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42682 MW;  575E7223D068CDDB CRC64;

  Query Match           100.0%;  Score 61;  DB 12;  Length 389;
  Best Local Similarity  100.0%;  Pred. No. 0.0016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db    149 PLSSIFSRIGDP 160


RESULT 10
Q9WMX3
ID   Q9WMX3      PRELIMINARY;      PRT;   389 AA.
AC   Q9WMX3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Surface antigen.
GN   S.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RX   MEDLINE=99238050; PubMed=10223539;
RA   Petzold D.R., Tautz B., Wolf F., Drescher J.;
RT   "Infection chains and evolution rates of Hepatitis B Virus in cardiac
RT   transplant recipients infected nosocommially.";
RL   J. Med. Virol. 58:1-10(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RA   Petzold D.R.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ131956; CAA10540.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
FT   CHAIN       164     389       SURFACE ANTIGEN.
SQ   SEQUENCE   389 AA;  42708 MW;  6AA08B85DFE5B99C CRC64;

  Query Match           100.0%;  Score 61;  DB 12;  Length 389;
```

```
  Best Local Similarity   100.0%;  Pred. No. 0.0016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 11
Q91C55
ID   Q91C55      PRELIMINARY;      PRT;   389 AA.
AC   Q91C55;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Large/middle/small S proteins.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=41;
RX   MEDLINE=21361467; PubMed=11468728;
RA   Owiredu W.K., Kramvis A., Kew M.C.;
RT   "Hepatitis B virus DNA in serum of healthy black African adults
RT   positive for hepatitis B surface antibody alone: possible association
RT   with recombination between genotypes A and D.";
RL   J. Med. Virol. 64:441-454(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=41;
RA   Owiredu W.K.B.A., Kramvis A., Kew M.C.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF297619; AAK97179.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42580 MW;  F4336E04992CE6E2 CRC64;

  Query Match             100.0%;  Score 61;  DB 12;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.0016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 12
Q91C52
ID   Q91C52      PRELIMINARY;      PRT;   389 AA.
AC   Q91C52;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Large/middle/small S proteins.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=57;
RX   MEDLINE=21361467; PubMed=11468728;
RA   Owiredu W.K., Kramvis A., Kew M.C.;
RT   "Hepatitis B virus DNA in serum of healthy black African adults
RT   positive for hepatitis B surface antibody alone: possible association
RT   with recombination between genotypes A and D.";
RL   J. Med. Virol. 64:441-454(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=57;
RA   Owiredu W.K.B.A., Kramvis A., Kew M.C.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF297620; AAK97183.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42937 MW;  1709DFD7D5E01066 CRC64;

  Query Match             100.0%;  Score 61;  DB 12;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.0016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 13
Q67875
ID   Q67875      PRELIMINARY;      PRT;   389 AA.
AC   Q67875;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Pre S1/S ORF.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lai M.E., Mazzoleni A.P., Balestrieri A., Melis A., Porru A.;
RT   "Sequence analysis of HBV genomes isolated from patients with HBsAg
RT   negative chronic liver disease.";
RL   Submitted (MAR-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X65258; CAA46353.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42661 MW;  A39542B416E48F24 CRC64;

  Query Match             100.0%;  Score 61;  DB 12;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.0016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 14
Q8QXP9
ID   Q8QXP9      PRELIMINARY;      PRT;   389 AA.
AC   Q8QXP9;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Large surface antigen.
GN   PRES1.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kay A.C.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Jeantet D., Chemin I., Mandrand B., Zoulim F., Trepo C., Kay A.;
RT   "Characterization of two Hepatitis B virus populations isolated from a
RT   Hepatitis B surface antigen-negative patient.";
RL   Hepatology 0:0-0(0).
DR   EMBL; AJ344117; CAC87013.1; -.
SQ   SEQUENCE   389 AA;  42764 MW;  9AC5A4D46BB73632 CRC64;

  Query Match             100.0%;  Score 61;  DB 12;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.0016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 PLSSIFSRIGDP 12
          ||||||||||||
Db    149 PLSSIFSRIGDP 160


RESULT 15
Q91IC3
ID   Q91IC3      PRELIMINARY;      PRT;    49 AA.
AC   Q91IC3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Pre-S2 protein (Fragment).
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=702-4;
RA   Huangfu J., Dong J., Deng H.;
RT   "Preliminary Study on PreS2 Region of Hepatitis B Virus.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF393221; AAK84360.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
FT   NON_TER      49     49
SQ   SEQUENCE   49 AA;  5045 MW;  5CA3FD6A356B59C3 CRC64;

  Query Match             96.7%;  Score 59;  DB 12;  Length 49;
  Best Local Similarity   91.7%;  Pred. No. 0.00041;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          |:||||||||||
Db     33 PISSIFSRIGDP 44


Search completed: February  3, 2003, 09:25:45
Job time : 2.94309 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:12 ; Search time 46.5488 Seconds
(without alignments)
1535.988 Million cell updates/sec

Title: US-09-890-752A-1
Perfect score: 1929
Sequence: 1 MGRGDGAGAFGLGFTPPHGG.....SILSPFLPLLPIFFCLWVYI 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1901	98.5	389	12	Q9WJC4	Q9wjc4 hepatitis b
2	1898	98.4	389	12	Q96843	Q96843 hepatitis b
3	1885	97.7	389	12	Q9PWW6	Q9pww6 hepatitis b
4	1883	97.6	389	12	Q67871	Q67871 hepatitis b
5	1883	97.6	389	12	Q67914	Q67914 hepatitis b
6	1879	97.4	389	12	Q67920	Q67920 hepatitis b
7	1878	97.4	389	12	Q9QMH6	Q9qmh6 hepatitis b
8	1876	97.3	389	12	Q9IF39	Q9if39 hepatitis b
9	1875	97.2	389	12	Q67893	Q67893 hepatitis b
10	1875	97.2	389	12	Q8QXQ3	Q8qxq3 hepatitis b
11	1875	97.2	389	12	Q8QXP9	Q8qxp9 hepatitis b
12	1874	97.1	389	12	Q9QAF2	Q9qaf2 hepatitis b
13	1864	96.6	389	12	Q9QAF9	Q9qaf9 hepatitis b
14	1864	96.6	389	12	Q9WNS1	Q9wns1 hepatitis b
15	1861	96.5	389	12	O92921	O92921 hepatitis b
16	1861	96.5	389	12	Q67879	Q67879 hepatitis b
17	1855	96.2	383	12	Q96839	Q96839 hepatitis b
18	1854	96.1	389	12	O90772	O90772 hepatitis b
19	1854	96.1	389	12	Q9WMX3	Q9wmx3 hepatitis b
20	1849	95.9	383	12	Q96838	Q96838 hepatitis b
21	1845	95.6	383	12	Q96837	Q96837 hepatitis b
22	1843	95.5	400	12	Q8V1J1	Q8v1j1 hepatitis b
23	1839	95.3	383	12	Q96840	Q96840 hepatitis b
24	1828	94.8	400	12	Q913A6	Q913a6 hepatitis b
25	1827	94.7	385	12	O92919	O92919 hepatitis b
26	1825	94.6	389	12	Q91C52	Q91c52 hepatitis b
27	1824	94.6	383	12	Q96836	Q96836 hepatitis b
28	1816	94.1	389	12	Q9QMI0	Q9qmi0 hepatitis b
29	1816	94.1	389	12	Q91C55	Q91c55 hepatitis b
30	1816	94.1	389	12	Q67890	Q67890 hepatitis b
31	1814	94.0	383	12	Q96841	Q96841 hepatitis b
32	1809.5	93.8	382	12	Q67886	Q67886 hepatitis b
33	1803	93.5	389	12	Q67956	Q67956 hepatitis b
34	1799	93.3	378	12	Q67954	Q67954 hepatitis b
35	1797	93.2	389	12	Q67875	Q67875 hepatitis b
36	1796	93.1	378	12	Q67944	Q67944 hepatitis b
37	1793	92.9	387	12	Q91A30	Q91a30 hepatitis b
38	1790	92.8	399	12	Q9QAW7	Q9qaw7 hepatitis b
39	1788	92.7	343	12	Q91T59	Q91t59 hepatitis b
40	1786	92.6	344	12	Q91TA6	Q91ta6 hepatitis b
41	1781	92.3	399	12	Q9IR25	Q9ir25 hepatitis b
42	1780	92.3	343	12	Q91T77	Q91t77 hepatitis b
43	1780	92.3	399	12	Q9IR27	Q9ir27 hepatitis b
44	1773	91.9	389	12	Q9QAW0	Q9qaw0 hepatitis b
45	1772	91.9	400	12	Q9E6T8	Q9e6t8 hepatitis b

ALIGNMENTS

RESULT 1
Q9WJC4
ID Q9WJC4 PRELIMINARY; PRT; 389 AA.
AC Q9WJC4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Pre-S and S proteins.
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85204397; PubMed=3996597;
RA Bichko V., Dreilina D., Pushko P.M., Pumpen P.P., Gren E.;
RT "Subtype ayw variant of hepatitis B virus.";
RL FEBS Lett. 185:208-212(1985).
DR EMBL; X02496; CAB41701.1; -.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; vMSA; 1.
SQ SEQUENCE 389 AA; 42745 MW; 6CD83C7CFD2BADF6 CRC64;

Query Match 98.5%; Score 1901; DB 12; Length 389;
Best Local Similarity 100.0%; Pred. No. 1.8e-154;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227
```

```
Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 2
Q96843
ID Q96843 PRELIMINARY; PRT; 389 AA.
AC Q96843;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE PreS protein.
GN PRES.
OS Hepatitis B virus (subtype ayw).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10418;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AYW;
RX MEDLINE=97332377; PubMed=9188611;
RA Stoll-Becker S., Repp R., Glebe D., Schaefer S., Kreuder J., Kann M.,
RA Lampert F., Gerlich W.H.;
RT "Transcription of hepatitis B virus in peripheral blood mononuclear
RT cells from persistently infected patients.";
RL J. Virol. 71:5399-5407(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AYW;
RA Repp R.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL; Y07587; CAA68861.1; -.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; vMSA; 1.
FT CHAIN 164 389 POTENTIAL.
SQ SEQUENCE 389 AA; 42744 MW; D8D4B1B9B92EE57B CRC64;

Query Match 98.4%; Score 1898; DB 12; Length 389;
Best Local Similarity 99.7%; Pred. No. 3.3e-154;
Matches 341; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 3
Q9PWW6
ID Q9PWW6 PRELIMINARY; PRT; 389 AA.
AC Q9PWW6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Large S protein.
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HBV/95-1758, AND HBV/94-11066;
RX MEDLINE=20109034; PubMed=10640544;
RA Hannoun C., Horal P., Lindh M.;
RT "Long-term mutation rates in the hepatitis B virus genome.";
RL J. Gen. Virol. 81:75-83(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HBV/95-1758, AND HBV/94-11066;
RA Hannoun C., Horal P., Lindh M.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF121241; AAF24674.1; -.
DR EMBL; AF121240; AAF24667.1; -.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; vMSA; 1.
SQ SEQUENCE 389 AA; 42814 MW; CD1898538A4A2C0C CRC64;

Query Match 97.7%; Score 1885; DB 12; Length 389;
Best Local Similarity 99.1%; Pred. No. 4.2e-153;
Matches 339; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||||:||| |||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPTNPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSHISSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 4
Q67871
ID Q67871 PRELIMINARY; PRT; 389 AA.
AC Q67871;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Pre S1/pre S2/S ORF.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.

```
RA   Lai M.E., Mazzoleni A.P., Balestrieri A., Melis A., Porru A.;
RT   "Sequence analysis of HBV genomes isolated from patients with HBsAg
RT   negative chronic liver disease.";
RL   Submitted (MAR-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X65257; CAA46349.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42756 MW;  977025C7DED114A1 CRC64;

  Query Match             97.6%;  Score 1883;  DB 12;  Length 389;
  Best Local Similarity   98.8%;  Pred. No. 6.2e-153;
  Matches  338;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||| | :||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASHLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 5
Q67914
ID   Q67914      PRELIMINARY;      PRT;   389 AA.
AC   Q67914;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Pre-S1 protein.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97025699; PubMed=8871878;
RA   Alexopoulou A., Karayiannis P., Hadziyannis S.J.;
RT   "Whole genome analysis of hepatitis B virus from 4 cases with
RT   fulminant hepatitis.";
RL   J. Viral Hepat. 3:173-181(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Karayiannis P.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X97848; CAA66425.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42717 MW;  495EF28627FBF628 CRC64;

  Query Match             97.6%;  Score 1883;  DB 12;  Length 389;
  Best Local Similarity   98.5%;  Pred. No. 6.2e-153;
  Matches  337;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGIIQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPVLNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSSHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 6
Q67920
ID   Q67920      PRELIMINARY;      PRT;   389 AA.
AC   Q67920;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Pre-S1 protein.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PRE-CORE VARIANT;
RX   MEDLINE=97025699; PubMed=8871878;
RA   Alexopoulou A., Karayiannis P., Hadziyannis S.J.;
RT   "Whole genome analysis of hepatitis B virus from 4 cases with
RT   fulminant hepatitis.";
RL   J. Viral Hepat. 3:173-181(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PRE-CORE VARIANT;
RA   Karayiannis P.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X97849; CAA66432.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42690 MW;  FD05FA0251E162F5 CRC64;

  Query Match             97.4%;  Score 1879;  DB 12;  Length 389;
  Best Local Similarity   98.5%;  Pred. No. 1.4e-152;
  Matches  337;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGIIQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSAGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347
```

```
Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 7
Q9QMH6
ID   Q9QMH6      PRELIMINARY;      PRT;   389 AA.
AC   Q9QMH6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   PreS1/preS2/S envelope.
GN   PRES1/PRES2/S.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GENOTYPE D;
RA   Okamoto H.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GENOTYPE D;
RX   MEDLINE=89010694; PubMed=3171552;
RA   Okamoto H., Tsuda F., Sakugawa H., Sastrosoewignjo R.I., Imai M.,
RA   Miyakawa Y., Mayumi M.;
RT   "Typing hepatitis B virus by homology in nucleotide sequence:
RT   comparison of surface antigen subtypes.";
RL   J. Gen. Virol. 69:2575-2583(1988).
DR   EMBL; AB033559; BAA85378.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42720 MW;  4D680B45D294FAEE CRC64;

  Query Match             97.4%;  Score 1878;  DB 12;  Length 389;
  Best Local Similarity   98.5%;  Pred. No. 1.7e-152;
  Matches  337;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||::|||||||||||||||||||||||||||| ||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGIIQTLPANPPPASTNRQSGRQPTPLSPPLRTTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCAGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 8
Q9IF39
ID   Q9IF39      PRELIMINARY;      PRT;   389 AA.
AC   Q9IF39;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   PreS1/preS2/S envelope.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HBV-D1;
RA   Yan L., Chen J.J., Hou J.L., Wang Z.H.;
RT   "First Chinese HBV strain: genotype D complete sequence report.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF280817; AAF81608.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42784 MW;  F53CCFC6A0C1F072 CRC64;

  Query Match             97.3%;  Score 1876;  DB 12;  Length 389;
  Best Local Similarity   98.5%;  Pred. No. 2.5e-152;
  Matches  337;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      6 GAGAPGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||:|||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db     48 GAGAYGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          ||||||| |||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    108 AMQWNSTIFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSHISSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFA 305
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFA 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 9
Q67893
ID   Q67893      PRELIMINARY;      PRT;   389 AA.
AC   Q67893;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Surface protein S.
GN   SURFACE PROTEIN S.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AYW4;
RA   Pluciennicczak A.;
RT   "Molecular cloning and sequencing of two complete genomes of polish
RT   isolates of Human Hepatitis B Virus.";
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z35716; CAA84788.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
FT   CHAIN       164    389       SURFACE PROTEIN S.
SQ   SEQUENCE   389 AA;  42762 MW;  5D27F0BA97A825D5 CRC64;

  Query Match             97.2%;  Score 1875;  DB 12;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 3e-152;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGIIQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSVGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTTVQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 10
Q8QXQ3
ID Q8QXQ3 PRELIMINARY; PRT; 389 AA.
AC Q8QXQ3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Large surface antigen.
GN PRES1.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HBSAG NEGATIVE PATIENT GX27;
RA Kay A.C.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HBSAG NEGATIVE PATIENT GX27;
RA Jeantet D., Chemin I., Mandrand B., Zoulim F., Trepo C., Kay A.;
RT "Characterization of two Hepatitis B virus populations isolated from a
RT Hepatitis B surface antigen-negative patient.";
RL Hepatology 0:0-0(0).
DR EMBL; AJ344116; CAC87019.1; -.
SQ SEQUENCE 389 AA; 42846 MW; E2A6E53DAD1353A6 CRC64;

Query Match 97.2%; Score 1875; DB 12; Length 389;
Best Local Similarity 98.5%; Pred. No. 3e-152;
Matches 337; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||| |||||||||||||||| |||||||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYLPAGGSSSGTVNPVPTTVSHISSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||:|| ||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYNILRPFLPLLPIFFCLWVYI 389
```

RESULT 11
Q8QXP9
ID Q8QXP9 PRELIMINARY; PRT; 389 AA.
AC Q8QXP9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Large surface antigen.
GN PRES1.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Kay A.C.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Jeantet D., Chemin I., Mandrand B., Zoulim F., Trepo C., Kay A.;
RT "Characterization of two Hepatitis B virus populations isolated from a
RT Hepatitis B surface antigen-negative patient.";
RL Hepatology 0:0-0(0).
DR EMBL; AJ344117; CAC87013.1; -.
SQ SEQUENCE 389 AA; 42764 MW; 9AC5A4D46BB73632 CRC64;

Query Match 97.2%; Score 1875; DB 12; Length 389;
Best Local Similarity 98.2%; Pred. No. 3e-152;
Matches 336; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:||||||||||||||:||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||||||| ||:|||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 12
Q9QAF2
ID Q9QAF2 PRELIMINARY; PRT; 389 AA.
AC Q9QAF2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Large S protein.
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

```
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HBV/1218;
RX   MEDLINE=20109034; PubMed=10640544;
RA   Hannoun C., Horal P., Lindh M.;
RT   "Long-term mutation rates in the hepatitis B virus genome.";
RL   J. Gen. Virol. 81:75-83(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HBV/1218;
RA   Hannoun C., Horal P., Lindh M.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF121242; AAF24681.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42910 MW;  BB6541E18D39357E CRC64;

  Query Match             97.1%;  Score 1874;  DB 12;  Length 389;
  Best Local Similarity   98.5%;  Pred. No. 3.7e-152;
  Matches  337;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           |||||||||||||||||||||||||||||:||| ||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPTNPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSHISSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILKIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           ||||||||||||||||||| ||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPRLYSILSPFLPLLPIFFCLWVYI 389


RESULT 13
Q9QAF9
ID   Q9QAF9      PRELIMINARY;      PRT;   389 AA.
AC   Q9QAF9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Large S protein.
GN   S.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HBV/1303;
RX   MEDLINE=20109034; PubMed=10640544;
RA   Hannoun C., Horal P., Lindh M.;
RT   "Long-term mutation rates in the hepatitis B virus genome.";
RL   J. Gen. Virol. 81:75-83(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HBV/1303;
RA   Hannoun C., Horal P., Lindh M.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF121239; AAF24661.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
SQ   SEQUENCE   389 AA;  42890 MW;  3D971AB01FD77FEC CRC64;

  Query Match             96.6%;  Score 1864;  DB 12;  Length 389;
  Best Local Similarity   98.5%;  Pred. No. 2.6e-151;
  Matches  337;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           |||||||||||||||||||||||||||||:||| ||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPTNPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSHISSIFSRIGDLALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMYLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           ||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 14
Q9WNS1
ID   Q9WNS1      PRELIMINARY;      PRT;   389 AA.
AC   Q9WNS1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Pre-S1 protein precursor.
GN   S.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AYW2;
RX   MEDLINE=20006001; PubMed=10534721;
RA   Gerner P., Lausch E., Friedt M., Tratzmuller R., Spangenberg C.,
RA   Wirth S.;
RT   "Hepatitis B virus core promoter mutations in children with multiple
RT   anti-HBe/HBeAg reactivations result in enhanced promoter activity.";
RL   J. Med. Virol. 59:415-423(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AYW2;
RA   Gerner P.R., Lausch E., Friedt M., Tratzmueller R., Spangenberg C.,
RA   Wirth S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF151735; AAD41361.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Signal.
FT   SIGNAL       1     54       POTENTIAL.
FT   CHAIN      164    389       S PROTEIN.
SQ   SEQUENCE   389 AA;  42779 MW;  30511E862904BE7C CRC64;

  Query Match             96.6%;  Score 1864;  DB 12;  Length 389;
  Best Local Similarity   98.0%;  Pred. No. 2.6e-151;
  Matches  335;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           |||||||||||||||||||||||||||||| |:||||||||||||||||||||||||||||||
```

```
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILHTVPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||:||| |||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISLIFSKIGDLALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db    288 TTPAQGTSMYPSCCCTKPSEGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||| ||||||||||||||||||||||||||||||||||||||
Db    348 GLSPIVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 15
O92921
ID   O92921      PRELIMINARY;      PRT;   389 AA.
AC   O92921;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Pre-S1 protein.
GN   S.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1/91;
RX   MEDLINE=98240943; PubMed=9581787;
RA   Gunther S., Paulij W., Meisel H., Will H.;
RT   "Analysis of hepatitis B virus populations in an interferon-alpha-
RT   treated patient reveals predominant mutations in the C-gene and
RT   changing e-antigenicity.";
RL   Virology 244:146-160(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1/91;
RA   Gunther S., Wilma P., Meisel H., Will H.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF043594; AAC40811.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
FT   CHAIN       164    389       S PROTEIN.
SQ   SEQUENCE   389 AA;  42795 MW;  3369787D78F188F9 CRC64;

  Query Match             96.5%;  Score 1861;  DB 12;  Length 389;
  Best Local Similarity   97.4%;  Pred. No. 4.7e-151;
  Matches  333;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||::||||||||||||||:||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGIIQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGLLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSVGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTYVQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||| ||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYRILSPFLPLLPIFFCLWVYI 389


Search completed: February  3, 2003, 09:25:43
Job time : 48.8821 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:16:26 ; Search time 12.6951 Seconds
(without alignments)
1133.686 Million cell updates/sec

Title: US-09-890-752A-1
Perfect score: 1929
Sequence: 1 MGRGDGAGAFGLGFTPPHGG.....SILSPFLPLLPIFFCLWVYI 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	1868	96.8	389	1	VMSA_HPBVA	P24025	hepatitis b
2	1868	96.8	389	1	VMSA_HPBVY	P03138	hepatitis b
3	1851	96.0	389	1	VMSA_HPBVZ	P03139	hepatitis b
4	1771	91.8	400	1	VMSA_HPBVR	P03140	hepatitis b
5	1755	91.0	389	1	VMSA_HPBVL	P12911	hepatitis b
6	1755	91.0	400	1	VMSA_HPBVP	Q02317	hepatitis b
7	1745	90.5	400	1	VMSA_HPBV4	P12934	hepatitis b
8	1715	88.9	389	1	VMSA_HPBVJ	P17398	hepatitis b
9	1713	88.8	400	1	VMSA_HPBV2	P03141	hepatitis b
10	1704	88.3	389	1	VMSA_HPBVO	P17399	hepatitis b
11	1695	87.9	389	1	VMSA_HPBVW	P03142	hepatitis b
12	1684	87.3	389	1	VMSA_HPBVI	P17397	hepatitis b
13	1684	87.3	400	1	VMSA_HPBV9	P17101	hepatitis b
14	1615	83.7	400	1	VMSA_HPBVT	Q05496	hepatitis b
15	1214	62.9	226	1	VMSA_HPBV0	P31873	hepatitis b
16	1211	62.8	226	1	VMSA_HPBVD	P31868	hepatitis b
17	1196	62.0	226	1	VMSA_HPBVN	P30019	hepatitis b
18	1188	61.6	226	1	VMSA_HPBVS	P31869	hepatitis b
19	938	48.6	431	1	VMSA_WHV59	P12910	woodchuck h
20	926	48.0	431	1	VMSA_WHV7	P12909	woodchuck h
21	924	47.9	431	1	VMSA_WHV8I	P17400	woodchuck h
22	922	47.8	428	1	VMSA_HPBGS	P03144	ground squi
23	918	47.6	426	1	VMSA_WHV1	P03143	woodchuck h
24	899	46.6	431	1	VMSA_WHV8	P06432	woodchuck h
25	833.5	43.2	282	1	VMSA_WHVW6	P11293	woodchuck h
26	220	11.4	335	1	VMSA_HPBHE	P13847	heron hepat
27	216.5	11.2	365	1	VMSA_HPBDC	P30029	duck hepati
28	216	11.2	366	1	VMSA_HPBDW	P17195	duck hepati
29	215	11.1	328	1	VMSA_HPBDU	P03145	duck hepati
30	213	11.0	366	1	VMSA_HPBDB	P17194	duck hepati
31	105	5.4	363	1	CYB_TRYBB	P00164	trypanosoma
32	105	5.4	503	1	WAIP_HUMAN	O43516	homo sapien
33	103.5	5.4	2504	1	FAS_HUMAN	P49327	homo sapien
34	103	5.3	1840	1	RPB1_ARATH	P18616	arabidopsis
35	100.5	5.2	1874	1	POLR_KYMVJ	P36304	kennedya ye
36	100	5.2	599	1	PGH1_HUMAN	P23219	homo sapien
37	99.5	5.2	1078	1	S24A_HUMAN	O95486	homo sapien
38	99.5	5.2	1887	1	RPB1_DROME	P04052	drosophila
39	98	5.1	531	1	TYD2_PAPSO	P54769	papaver som
40	97.5	5.1	759	1	TEA1_YEAST	P47988	saccharomyc
41	97.5	5.1	2359	1	CCAH_RAT	Q9eq60	rattus norv
42	97	5.0	605	1	YHC8_YEAST	P38739	saccharomyc
43	96	5.0	371	1	CYB_LEITA	P14548	leishmania
44	96	5.0	626	1	GPBA_HUMAN	P07359	homo sapien
45	95.5	5.0	378	1	LEUK_RAT	P13838	rattus norv

ALIGNMENTS

RESULT 1
VMSA_HPBVA
ID VMSA_HPBVA STANDARD; PRT; 389 AA.
AC P24025;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (strain alpha1).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10411;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90266476; PubMed=2345966;
RA Tong S., Li J., Vitvitski L., Trepo C.;
RT "Active hepatitis B virus replication in the presence of anti-HBe is
RT associated with viral variants containing an inactive pre-C region.";
RL Virology 176:596-603(1990).
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; M32138; AAA45502.1; -.
DR PIR; B34773; SAVLA1.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; vMSA; 1.
KW Antigen.
FT PROPEP 1 163
FT CHAIN 164 389 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 389 AA; 42733 MW; B806D0D42B223BF5 CRC64;

Query Match 96.8%; Score 1868; DB 1; Length 389;
Best Local Similarity 97.7%; Pred. No. 2.3e-125;
Matches 334; Conservative 5; Mismatches 3; Indels 0; Gap

Qy 6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHP
Db 48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLR

Qy 66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIG
Db 108 AMQWNSTTFHQTLQDPRVRGLYLPAGGSSSGTVNPVPTTVSPISSIFS

Qy 126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVC
Db 168 TSGFLGPLLVLQAGFFLLTKILTIPKSLDSWWTSLNFLGGT

page 2

```
Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||| |||||||||| |::|||||||||||||||||||
Db    348 GLSPTVWLLVIWMMWYWGPKLFTILSPFLPLLPIFFCLWVYI 389
```

RESULT 2
VMSA_HPBVY
ID VMSA_HPBVY STANDARD; PRT; 389 AA.
AC P03138;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (subtype ayw).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10418;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81012091; PubMed=399327;
RA Galibert F., Mandart E., Fitoussi F., Tiollais P., Charnay P.;
RT "Nucleotide sequence of the hepatitis B virus genome (subtype ayw)
RT cloned in E. coli.";
RL Nature 281:646-650(1979).
RN [2]
RP SEQUENCE FROM N.A. (CLONE PHB320).
RX MEDLINE=85204397; PubMed=3996597;
RA Bichko V., Pushko P., Dreilina D., Pumpen P., Gren E.Y.;
RT "Subtype ayw variant of hepatitis B virus. DNA primary structure
RT analysis.";
RL FEBS Lett. 185:208-212(1985).
RN [3]
RP SEQUENCE FROM N.A.
RA Pumpen P., Kozlovskaya T.M., Borisova G.L., Byuchko V.V.,
RA Dishler A.V., Kalic Y.V., Pudova N.V., Gren E.Y., Krievinya V.Y.,
RA Kukain R.A.;
RT "Synthesis of the surface antigen of hepatitis B virus in Escherichia
RT coli.";
RL Dokl. Biochem. 271:246-249(1984).
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; V01460; -; NOT_ANNOTATED_CDS.
DR EMBL; X02496; CAA26324.1; ALT_INIT.
DR EMBL; M12393; AAA45496.1; ALT_INIT.
DR PIR; A03703; SAVLAH.
DR PIR; A03704; SAVLBH.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; vMSA; 1.
KW Antigen.
FT PROPEP 1 163
FT CHAIN 164 389 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 75 75 Q -> E (IN REF. 2).
FT CONFLICT 147 147 A -> S (IN REF. 2).
FT CONFLICT 150 150 L -> I (IN REF. 2).
FT CONFLICT 288 290 MTT -> TTP (IN REF. 2 AND 3).
SQ SEQUENCE 389 AA; 42766 MW; 6DC9E682DA694F63 CRC64;

```
  Query Match             96.8%;  Score 1868;  DB 1;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 2.3e-125;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 3
VMSA_HPBVZ
ID VMSA_HPBVZ STANDARD; PRT; 389 AA.
AC P03139;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (subtype adyw).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10419;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81012115; PubMed=399329;
RA Pasek M., Goto T., Gilbert W., Zink B., Schaller H., Mackay P.,
RA Leadbetter G., Murray K.;
RT "Hepatitis B virus genes and their expression in E. coli.";
RL Nature 282:575-579(1979).
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; J02202; AAA45487.1; ALT_INIT.
DR PIR; A93217; SAVLAJ.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; vMSA; 1.
KW Antigen.
FT PROPEP 1 163
FT CHAIN 164 389 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 389 AA; 42801 MW; 1069EF6B4679A669 CRC64;

```
  Query Match             96.0%;  Score 1851;  DB 1;  Length 389;
  Best Local Similarity   97.1%;  Pred. No. 3.7e-124;
  Matches  332;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||||::||||||||||||||||||||||||| ||||
Db      48 GAGAPGLGFTPPHGGLLGWSPQAQGIMQTLPANPPPASTNRQSGRQPTPLSPPLRTTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          || |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     108 AMHWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTTSPISSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPISNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGSCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTPAQGISMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||| |||||||||||||||||||||||||||||| ||
Db     348 GLSPIVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWAYI 389


RESULT 4
VMSA_HPBVR
ID   VMSA_HPBVR     STANDARD;      PRT;   400 AA.
AC   P03140;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adr).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=106820;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83168919; PubMed=6300776;
RA   Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
RT   "The complete nucleotide sequences of the cloned hepatitis B virus
RT   DNA; subtype adr and adw.";
RL   Nucleic Acids Res. 11:1747-1757(1983).
RN   [2]
RP   SEQUENCE OF 176-350 FROM N.A.
RX   MEDLINE=85107103; PubMed=3968537;
RA   Takeshima H., Inokoshi J., Namiki M., Shimada J., Omura S.;
RT   "Structural analysis of the gene coding for hepatitis B virus surface
RT   antigen and its product.";
RL   J. Gen. Virol. 66:195-200(1985).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00867; CAA24234.1; ALT_INIT.
DR   PIR; A03705; SAVLA.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    174
FT   CHAIN       175    400       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD     15     15       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    123    123       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   400 AA;  43685 MW;  16DB0B15C6CB697A CRC64;

  Query Match             91.8%;  Score 1771;  DB 1;  Length 400;
  Best Local Similarity   92.1%;  Pred. No. 1.7e-118;
  Matches  315;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||| |||||||||||||||||||| |:|| ||||||||||||||||:|||||::|||
Db      59 GAGAFGPGFTPPHGGLLGWSPQAQGILTTVPAAPPPASTNRQSGRQPTPISPPLRDSHPQ 118

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||| | |||||||||||||||||||||||||| ||||||||| |||| ||||
Db     119 AMQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRTGDPAPNMENT 178

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||    | ||||||||||||||||
Db     179 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 238

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||| |||||||||||||||||||||||||||||||||||||||||||:||:|||||||:||
Db     239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 298

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | |||||||:|||||||||||||||||||||||| :|||||||  ||||||||||||||||
Db     299 TIPAQGTSMFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFV 358

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||:|||||||||||||||||||||
Db     359 GLSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 400


RESULT 5
VMSA_HPBVL
ID   VMSA_HPBVL     STANDARD;      PRT;   389 AA.
AC   P12911;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-APR-1990 (Rel. 14, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (strain lsh / chimpanzee isolate).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10414;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88258473; PubMed=2838576;
RA   Vaudin M., Wolstenholme A.J., Tsiquaye K.N., Zuckerman A.J.,
RA   Harrison T.J.;
RT   "The complete nucleotide sequence of the genome of a hepatitis B
RT   virus isolated from a naturally infected chimpanzee.";
RL   J. Gen. Virol. 69:1383-1389(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00220; BAA00159.1; -.
DR   PIR; C28885; SAVLCP.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    163
FT   CHAIN       164    389       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD    309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   389 AA;  42539 MW;  87D1F93382A10DBD CRC64;

  Query Match             91.0%;  Score 1755;  DB 1;  Length 389;
  Best Local Similarity   91.5%;  Pred. No. 2.3e-117;
  Matches  313;  Conservative    9;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           |||||| |||||||||||| |||||||| |||||||||||||||||||||||||:|||| 
Db      48 GAGAFGPGFTPPHGGLLGLSPQAQGILTTLPANPPPASTNRQSGRQPTPLSPPLRDTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||| |||||||||||||||||:|||| | | |||:||  |||| |||||
Db     108 AMQWNSTTFHQALQDPRVRGLYFPAGGSSSGTLNPVPNTASHISSVFSTTGDPAPNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||:|||||||||||||||||  ||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTKILTIPQSLDSWWTSLNFLGGAPVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db     228 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCKTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           ||||||||: |||||||||||||||||||||||||| |||||||| |||||||| ||||||
Db     288 TTPAQGTSLIPSCCCTKPSDGNCTCIPIPSSWAFAKFLWEWASVRFSWLSLLAPFVQWFA 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           ||||||||  |||||||||:||:||||||:|||||||||||||
Db     348 GLSPTVWLLAIWMMWYWGPNLYNILSPFIPLLPIFFCLWVYI 389


RESULT 6
VMSA_HPBVP
ID   VMSA_HPBVP     STANDARD;      PRT;   400 AA.
AC   Q02317;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adw / strain Philippino/pFDW294).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=31514;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Estacio R.C., Chavez C.C., Okamoto H., Lingao A.L., Reyes M.T.,
RA   Domingo E., Mayumi M.;
RT   "Nucleotide sequence of a hepatitis B virus genome of subtype adw
RT   isolated from a Philippino: comparison with the reported three genomes
RT   of the same subtype.";
RL   J. Gastroenterol. Hepatol. 3:215-222(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57663; AAA69680.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    174
FT   CHAIN       175    400       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD     15     15       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    123    123       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   400 AA;  43740 MW;  7EBD53A004136BD2 CRC64;

  Query Match             91.0%;  Score 1755;  DB 1;  Length 400;
  Best Local Similarity   90.9%;  Pred. No. 2.4e-117;
  Matches  311;  Conservative   13;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           | |||| ||||||||:|||||||||||| |:|| |||||||||||||||:|||||::|||
Db      59 GVGAFGPGFTPPHGGVLGWSPQAQGILATVPAMPPPASTNRQSGRQPTPISPPLRDSHPQ 118

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||  ||| |||||||||||||||||:|||||   | |||| |||||| |||||
Db     119 AMQWNSTAFHQALQDPRVRGLYFPAGGSSSGTLNPVPTIASHISSISSRIGDPAPNMENI 178

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db     179 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPVCLGQNSQSPTSNHSPTS 238

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| ||||||||||||||||||||||||||||||||||||||||||||:||||||:||
Db     239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSTTTSTGPCKTC 298

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           |||||| ||:||||||||:|||||||||||||| |:|||||| ||||||||||||||||||
Db     299 TTPAQGNSMPPSCCCTKPTDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFV 358

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||| |||||||||:|||||||||:|||||||||||||||
Db     359 GLSPTVWLSAIWMMWYWGPSLYNILSPFIPLLPIFFCLWVYI 400


RESULT 7
VMSA_HPBV4
ID   VMSA_HPBV4     STANDARD;      PRT;   400 AA.
AC   P12934; Q67860;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adr4).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10409;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83246570; PubMed=6306594;
RA   Fujiyama A., Miyanohara A., Nozaki C., Yoneyama T., Ohtomo N.,
RA   Matsubara K.;
RT   "Cloning and structural analyses of hepatitis B virus DNAs, subtype
RT   adr.";
RL   Nucleic Acids Res. 11:4601-4610(1983).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X01587; CAA25747.1; ALT_INIT.
DR   EMBL; X01587; CAA25743.1; ALT_INIT.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    174
FT   CHAIN       175    400       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD     15     15       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    123    123       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   400 AA;  43542 MW;  01053BBCD24053F9 CRC64;

  Query Match             90.5%;  Score 1745;  DB 1;  Length 400;
  Best Local Similarity   90.9%;  Pred. No. 1.2e-116;
```

```
  Matches  311;  Conservative   11;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGPTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           |||||| |||||||||||||||||:| |:|  |||||||||||||||||:|||||::|||
Db      59 GAGAFGPGFTPPHGGLLGWSPQAQGVLTTVPVAPPPASTNRQSGRQPTPISPPLRDSHPQ 118

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||  |||||||||||||||||||||||||||| |||||  || |||| ||||
Db     119 AMQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTASPISSISSRTGDPAPNMENT 178

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           |||||||||||||||||||||||||||||||||||||||||||    | ||||||||||||||
Db     179 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 238

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| |||||||||||||||||||||||||||||||||||||||||:||:|||||||:||
Db     239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 298

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||:|||||||||||||||||||||||| :|||| || ||||||||||||||||
Db     299 TIPAQGTSMFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEGASVRFSWLSLLVPFVQWFV 358

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||:||||||||||||||||||||||
Db     359 GLSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 400


RESULT 8
VMSA_HPBVJ
ID   VMSA_HPBVJ     STANDARD;      PRT;   389 AA.
AC   P17398;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adw / strain Japan/pJDW233).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10413;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89010694; PubMed=3171552;
RA   Okamoto H., Tsuda F., Sakugawa H., Sastrosoewignjo R.I., Imai M.,
RA   Miyakawa Y., Mayumi M.;
RT   "Typing hepatitis B virus by homology in nucleotide sequence:
RT   comparison of surface antigen subtypes.";
RL   J. Gen. Virol. 69:2575-2583(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00329; -; NOT_ANNOTATED_CDS.
DR   PIR; G28925; SAVLJ1.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    163
FT   CHAIN       164    389       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   389 AA;  42603 MW;  72F8D481924DC104 CRC64;

  Query Match             88.9%;  Score 1715;  DB 1;  Length 389;
  Best Local Similarity   89.5%;  Pred. No. 1.5e-114;
  Matches  306;  Conservative   12;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           | |||| |||||||||||||||||||| ::|| |||||||||||||||||||||:||||
Db      48 GVGAFGPGFTPPHGGLLGWSPQAQGILTSVPAAPPPASTNRQSGRQPTPLSPPLRDTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||| |||||||||||:|   ||| |||| |: |||  |||||
Db     108 AMQWNSTTFHQTLQDPRVRALYFPAGGSSSGTVSPAQNTVSAISSILSKTGDPVPNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           || |||||||||||||:|||||||||||||||||||| |||||||||  |:||||
Db     168 ASGLLGPLLVLQAGFFLLTKILTIPQSLDSWWTSLNFLGGTPVCLGQNSQSQISSHSPTC 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| |||||||||||||| |||||||||||||||||||||||||||||||||||:||
Db     228 CPPICPGYRWMCLRRFIIFLCILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCKTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           |||||||||:|||||||| |||||||||||||| |:|||||| |||||||||||||||||
Db     288 TTPAQGTSMFPSCCCTKPMDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||:|||||:||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYNILSPFMPLLPIFFCLWVYI 389


RESULT 9
VMSA_HPBV2
ID   VMSA_HPBV2     STANDARD;      PRT;   400 AA.
AC   P03141;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adw2).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10408;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Valenzuela P., Quiroga M., Zaldivar J., Gray P., Rutter W.J.;
RL   (In) Field B.N., Jaenisch R., Fox C.F. (eds.);
RL   Animal virus genetics, pp.57-70, Academic Press, New York (1980).
RN   [2]
RP   SEQUENCE OF 175-400 FROM N.A.
RX   MEDLINE=79244739; PubMed=471053;
RA   Valenzuela P., Gray P., Quiroga M., Zaldivar J., Goodman H.M.,
RA   Rutter W.J.;
RT   "Nucleotide sequence of the gene coding for the major protein of
RT   hepatitis B virus surface antigen.";
RL   Nature 280:815-819(1979).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X02763; CAA26539.1; -.
DR   EMBL; J02205; AAA45524.1; -.
DR   PIR; A03706; SAVLVD.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    174
FT   CHAIN       175    400       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD     15     15       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    123    123       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   400 AA;  43704 MW;  57356B6293872BC5 CRC64;
```

```
  Query Match             88.8%;  Score 1713;  DB 1;  Length 400;
  Best Local Similarity   88.6%;  Pred. No. 2.2e-114;
  Matches  303;  Conservative   14;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           | ||||    |||||||:||||||||||| |:   |||||||||||||||||:|||||::|||
Db      59 GVGAFGPRLTPPHGGILGWSPQAQGILTTVSTIPPPASTNRQSGRQPTPISPPLRDSHPQ 118

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           ||||||| |||||||||||||| |||||||||||| |    | |||| :| |||  |||||
Db     119 AMQWNSTAFHQTLQDPRVRGLYLPAGGSSSGTVNPAPNIASHISSISARTGDPVTNMENI 178

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||: ||||||||||||||||
Db     179 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGSPVCLGQNSQSPTSNHSPTS 238

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| |||||||||||||||||||||||||||||||||||||||||||||:|||||||:||
Db     239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSTTTSTGPCKTC 298

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           |||||| ||:||||||||:|||||||||||||||| |:|||||| |||||||||||||||
Db     299 TTPAQGNSMFPSCCCTKPTDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFV 358

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||| |||||||||||||:|||:|||||||||||||
Db     359 GLSPTVWLSAIWMMWYWGPSLYSIVSPFIPLLPIFFCLWVYI 400


RESULT 10
VMSA_HPBVO
ID   VMSA_HPBVO     STANDARD;      PRT;   389 AA.
AC   P17399;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adw / strain Okinawa/pODW282).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10415;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89010694; PubMed=3171552;
RA   Okamoto H., Tsuda F., Sakugawa H., Sastrosoewignjo R.I., Imai M.,
RA   Miyakawa Y., Mayumi M.;
RT   "Typing hepatitis B virus by homology in nucleotide sequence:
RT   comparison of surface antigen subtypes.";
RL   J. Gen. Virol. 69:2575-2583(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00330; -; NOT_ANNOTATED_CDS.
DR   PIR; H28925; SAVLJ2.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    163
FT   CHAIN       164    389       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   389 AA;  42408 MW;  CECAC0DDA3D84A10 CRC64;

  Query Match             88.3%;  Score 1704;  DB 1;  Length 389;
  Best Local Similarity   88.6%;  Pred. No. 9.3e-114;
  Matches  303;  Conservative   15;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           | ||||  |||||||||||||||||||:| |:|| |||||||||||||||||||||:||||
Db      48 GVGAFGPGFTPPHGGLLGWSPQAQGLLTTVPAAPPPASTNRQSGRQPTPLSPPLRDTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||| || |||||||||||||:|   |||| |||| |: |||  |||||
Db     108 AMQWNSTTFHQTLQDPGVRALYFPAGGSSSGTVSPAQNTVSAISSILSKTGDPVPNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           || |||||||||||||:|||||||||||||||||||||| |||||||||  |:||||
Db     168 ASGLLGPLLVLQAGFFLLTKILTIPQSLDSWWTSLNFLGGTPVCLGQNSQSQISSHSPTC 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| |||||||||||||| ||||||||||||||||||||||||||||||||||||:||
Db     228 CPPICPGYRWMCLRRFIIFLCILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCKTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           |||||||||:||||||||:|||||||||||||||| |:|||||| |||||||||||||||
Db     288 TTPAQGTSMFPSCCCTKPTDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFMPLLPIFFCLWVYI 347
           |||||||||||||:|:||||||:|||||:||||||||||||
Db     348 GLSPTVWLSVIWMIWFWGPSLYNILSPFMPLLPIFFCLWVYI 389


RESULT 11
VMSA_HPBVW
ID   VMSA_HPBVW     STANDARD;      PRT;   389 AA.
AC   P03142;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adw).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=106821;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83168919; PubMed=6300776;
RA   Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
RT   "The complete nucleotide sequences of the cloned hepatitis B virus
RT   DNA; subtype adr and adw.";
RL   Nucleic Acids Res. 11:1747-1757(1983).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00866; CAA24233.1; ALT_INIT.
DR   PIR; A93460; SAVLVE.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    163
FT   CHAIN       164    389       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   389 AA;  42354 MW;  FE9EF2E7EEFC58B5 CRC64;

  Query Match             87.9%;  Score 1695;  DB 1;  Length 389;
  Best Local Similarity   87.7%;  Pred. No. 4e-113;
  Matches  300;  Conservative   14;  Mismatches   28;  Indels    0;  Gaps    0;
```

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
Db     48 GVGAFGPGLTPPHGGILGWSPQAQGILTTVSTIPPPASTNRQSGRQPTPISPPLRDSHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
Db    108 AMQWNSTALHQALQDPRVRGLYLPAGGSSSGTVNPAPNIASHISSISARTGDPVTIMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGSPVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
Db    228 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSTTTSTGPCKTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
Db    288 TTPAQGNSKFPSCCCTKPTDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
Db    348 GLSPTVWLSAIWMMWYWGPSLYSIVSPFIPLLPIFFCLWVYI 389
```

RESULT 12
VMSA_HPBVI
ID VMSA_HPBVI STANDARD; PRT; 389 AA.
AC P17397;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (subtype adw / strain Indonesia/pIDW420).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10412;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89010694; PubMed=3171552;
RA Okamoto H., Tsuda F., Sakugawa H., Sastrosoewignjo R.I., Imai M.,
RA Miyakawa Y., Mayumi M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes.";
RL J. Gen. Virol. 69:2575-2583(1988).
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; D00331; -; NOT_ANNOTATED_CDS.
DR PIR; I28925; SAVLJ3.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; vMSA; 1.
KW Antigen.
FT PROPEP 1 163
FT CHAIN 164 389 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 389 AA; 42653 MW; 08079BA34F3B90C0 CRC64;

Query Match 87.3%; Score 1684; DB 1; Length 389;
Best Local Similarity 88.0%; Pred. No. 2.4e-112;
Matches 301; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
Db     48 GVGAFGPGFTPPHGGLLGWSPQAQGILTTVPTAPPPASTNRQLGRKPTPLSPPLRDTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
Db    108 AMQWNSTTFHQTLQDPRVRALYFPAGGSSSGTVNPVQNTASSISSILSTTGDPVPNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
Db    168 ASGLLGPLLVLQAGFFSLTKILTIPLSLDSWWTSLNFLGETPVCLGQNSQSQISSHSPTC 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
Db    228 CPPICPGYRWMCLRRFIIFLCILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCKTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
Db    288 TTPAQGTSMFPSCCCTKPTDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
Db    348 GLSPTVWLSVIWMMWFWGPSLYNILSPFMPLLPIFFCLWVYI 389
```

RESULT 13
VMSA_HPBV9
ID VMSA_HPBV9 STANDARD; PRT; 400 AA.
AC P17101;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (subtype adw / strain 991).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10410;
RN [1]
RP SEQUENCE FROM N.A.
RA Koechel H.G., Schueler A., Lottmann S., Thomssen R.;
RL Submitted (FEB-1990) to the EMBL/GenBank/DDBJ databases.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; X51970; CAA36230.1; -.
DR PIR; S10383; SAVLKS.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; vMSA; 1.
KW Antigen.
FT PROPEP 1 174
FT CHAIN 175 400 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 400 AA; 43695 MW; 76D98F74B8B17724 CRC64;

Query Match 87.3%; Score 1684; DB 1; Length 400;
Best Local Similarity 87.1%; Pred. No. 2.5e-112;
Matches 298; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
Db     59 GVGAFGPGFTPPHGGVLGWSPQAQGMLTPVSTIPPPASANRQSGRQPTPISPPLRDSHPQ 118

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
Db    119 AMQWNSTAFHQALQDPRVRGLYFPAGGSSSGTVNPAPNIASHISSISARTGDPVTNMENI 178
```

```
Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||| |||||||||||||||||||||||||||||||||: |||||||:|||||||||||
Db    179 TSGFLGPLPVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGSPVCLGQNSRSPTSNHSPTS 238

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||| ||||||||||||||||||||||||||||||||||||||||||||| ||:|||||||:||
Db    239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLILGSTTTSTGPCKTC 298

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          |||||| ||:|||||||:|||||||||||||| |:|||||| ||||||||||||||||||
Db    299 TTPAQGNSMPPSCCCTKPTDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFV 358

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          ||||||||| |||||||||||||:| |:|||||||||||||
Db    359 GLSPTVWLSAIWMMWYWGPSLYSIVSSFIPLLPIFFCLWVYI 400


RESULT 14
VMSA_HPBVT
ID   VMSA_HPBVT     STANDARD;      PRT;   400 AA.
AC   Q05496;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adw4 / strain Brazil / isolate w4B).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=45410;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93346970; PubMed=8345355;
RA   Naumann H., Schaefer S., Yoshida C.F.T., Gaspar A.M.C., Repp R.,
RA   Gerlich W.H.;
RT   "Identification of a new hepatitis B virus (HBV) genotype from Brazil
RT   that expresses HBV surface antigen subtype adw4.";
RL   J. Gen. Virol. 74:1627-1632(1993).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X69798; CAA49455.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    174
FT   CHAIN       175    400       MAJOR SURFACE ANTIGEN.
SQ   SEQUENCE   400 AA;  43551 MW;  4A5A2212E4B3E117 CRC64;

  Query Match             83.7%;  Score 1615;  DB 1;  Length 400;
  Best Local Similarity   84.5%;  Pred. No. 1.9e-107;
  Matches  289;  Conservative   20;  Mismatches   33;  Indels    0;  Gaps    0;

Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          | | :| |||||||||||||||||:| ||||:||||||||:|||:|||:|||||:||||
Db     59 GVGGYGPGFTPPHGGLLGWSPQAQGVLTTLPADPPPASTNRRSGRKPTPVSPPLRDTHPQ 118

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          ||||||| ||| | ||||| ||||||||||||| || ||  |  |||||: | ||:||:||
Db    119 AMQWNSTQFHQALLDPRVRALYFPAGGSSSGTQNPAPTIASLTSSIFSKTGGPAMNMDNI 178

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||| |||||||||  ||||:|||||||||||||||||    | ||||||||| |||
Db    179 TSGLLGPLLVLQAVCPLLTKILTIPQSLDSWWTSLNFLGGLPGCPGQNSQSPTSNHLPTS 238

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||:||
Db    239 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGSTTTSTGPCKTC 298

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          || ||||||:|||||:|||||||||||||||| ||:|||||||||||||||||| |||| |
Db    299 TTLAQGTSMFPSCCCSKPSDGNCTCIPIPSSWALGKYLWEWASARFSWLSLLVQFVQWCV 358

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||| ||||:|||||:| ||||||:||||||  ||| |
Db    359 GLSPTVWLLVIWMIWYWGPNLCSILSPFIPLLPIFCYLWVSI 400


RESULT 15
VMSA_HPBV0
ID   VMSA_HPBV0     STANDARD;      PRT;   226 AA.
AC   P31873;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   Major surface antigen.
GN   S.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88297159; PubMed=2841200;
RA   Rivkina M.B., Lunin V.G., Mahov A.M., Tikchonenko T.I., Kukain R.A.;
RT   "Nucleotide sequence of integrated hepatitis B virus DNA and human
RT   flanking regions in the genome of the PLC/PRF/5 cell line.";
RL   Gene 64:285-296(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21030; AAA45516.1; -.
DR   EMBL; X04820; CAA28506.1; -.
DR   PIR; JT0293; SAVLHV.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   CARBOHYD      3      3       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    146    146       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   226 AA;  25476 MW;  BFC4329CF1720600 CRC64;

  Query Match             62.9%;  Score 1214;  DB 1;  Length 226;
  Best Local Similarity   93.4%;  Pred. No. 2.5e-79;
  Matches  211;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    122 MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNH 181
          |||||||||||||||||||||||||||||||||||||||||||||: |||||||||||||||
Db      1 MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGSPVCLGQNSQSPTSNH 60

Qy    182 SPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGP 241
          ||||||| |||||||||||||||||||||||||||||||||||||||||||||:||||||
Db     61 SPTSCPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSTTTSTGP 120

Qy    242 CRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFV 301
          |:|||||||| ||:||||||:|||||||||||||| |:|||| | ||||||||||||
Db    121 CKTCTTPAQGNSMFPSCCCTKPTDGNCTCIPIPSSWAFAKYLWEWGSVRFSWLSLLVPFV 180

Qy    302 QWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||:||:||||||:||||||||||||
Db    181 QWFVGLSPTVWLSVIWMMWYWGPNLYNILSPFIPLLPIFFCLWVYI 226


Search completed: February  3, 2003, 09:24:18
Job time : 14.6951 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:57 ; Search time 22.9721 Seconds
(without alignments)
1452.136 Million cell updates/sec

Title: US-09-890-752A-1
Perfect score: 1929
Sequence: 1 MGRGDGAGAFGLGFTPPHGG.....SILSPFLPLLPIFFCLWVYI 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1901	98.5	389	1	SAVLBH	large surface anti
2	1883	97.6	389	2	S20745	surface antigen -
3	1879	97.4	389	2	S32202	large surface anti
4	1875	97.2	389	2	S47407	surface antigen -
5	1868	96.8	389	1	SAVLAH	large surface anti
6	1868	96.8	389	1	SAVLA1	large surface anti
7	1861	96.5	389	2	S20753	surface antigen -
8	1851	96.0	389	1	SAVLAJ	large surface anti
9	1803	93.5	389	2	S41871	surface antigen -
10	1799	93.3	378	2	S41869	surface antigen -
11	1797	93.2	389	2	S20749	surface antigen -
12	1771	91.8	400	1	SAVLA	large surface anti
13	1755	91.0	389	1	SAVLCP	large surface anti
14	1745	90.5	400	2	S35528	surface antigen -
15	1734	89.9	378	2	S41870	surface antigen -
16	1723	89.3	389	2	S67506	large surface anti
17	1715	88.9	389	1	SAVLJ1	large surface anti
18	1713	88.8	400	1	SAVLVD	large surface anti
19	1713	88.8	400	1	JQ1575	major surface anti
20	1711	88.7	445	2	S43492	surface antigen -
21	1704	88.3	389	1	SAVLJ2	large surface anti
22	1699	88.1	389	2	S36654	surface antigen -
23	1695	87.9	389	1	SAVLVE	large surface anti
24	1684	87.3	389	1	SAVLJ3	large surface anti
25	1684	87.3	400	1	SAVLKS	large surface anti
26	1641	85.1	384	2	T13469	large surface anti
27	1615	83.7	400	2	JQ2230	large surface prot
28	1593	82.6	382	2	T13474	large surface anti
29	1324	68.6	281	2	JQ2226	middle surface pro
30	1269	65.8	226	2	JQ2077	surface antigen -
31	1269	65.8	226	2	JQ2076	surface antigen -
32	1267	65.7	226	1	JQ1571	major surface anti
33	1265	65.6	226	2	JQ2075	surface antigen -
34	1264	65.5	226	2	JQ2079	surface antigen -
35	1264	65.5	226	2	JQ2078	surface antigen -
36	1263	65.5	226	2	JQ2069	surface antigen -
37	1260	65.3	226	2	JQ2073	surface antigen -
38	1259	65.3	226	2	JQ2067	surface antigen -
39	1257	65.2	226	2	JQ2083	surface antigen -
40	1256	65.1	226	2	JQ2063	surface antigen -
41	1253	65.0	226	1	JQ1572	major surface anti
42	1253	65.0	226	2	JQ2065	surface antigen -
43	1250	64.8	226	2	JQ2070	surface antigen -
44	1250	64.8	226	2	JQ2066	surface antigen -
45	1249	64.7	226	2	JQ2081	surface antigen -

ALIGNMENTS

RESULT 1
SAVLBH
large surface antigen - hepatitis B virus (subtype ayw, strain pHB320)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
A;Note: host Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 28-Jul-2000
C;Accession: A03704; PQ0585
R;Bichko, V.; Pushko, P.; Dreilina, D.; Pumpen, P.; Gren, E.
FEBS Lett. 185, 208-212, 1985
A;Title: Subtype ayw variant of hepatitis B virus: DNA primary structure analysis.
A;Reference number: A05237; MUID:85204397; PMID:3996597
A;Accession: A03704
A;Molecule type: DNA
A;Residues: 1-389 <BIC>
A;Cross-references: GB:X02496; NID:g62280; PIDN:CAB41701.1; PID:g4704321
R;Norder, H.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 3141-3145, 1992
A;Title: Molecular basis of hepatitis B virus serotype variations within the four major
A;Reference number: PQ0453; MUID:93107848; PMID:1469353
A;Accession: PQ0585
A;Molecule type: DNA
A;Residues: 264-343 <NOR>
A;Experimental source: subtype ayw2, Tav
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 98.5%; Score 1901; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 3.3e-128;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287
```

```
Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 2
S20745
surface antigen - hepatitis B virus (subtype ayw, patient C)
N;Alternate names: HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient C
C;Date: 20-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
C;Accession: S20745
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submitted to the EMBL Data Library, March 1992
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati
A;Reference number: S20745
A;Accession: S20745
A;Molecule type: DNA
A;Residues: 1-389 <LAI>
A;Cross-references: EMBL:X65257; NID:g59429; PIDN:CAA46349.1; PID:g59430
A;Experimental source: subtype ayw, patient C
C;Genetics:
A;Gene: S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-108/Domain: pre-S1 domain #status predicted <PRE1>
F;109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
F;109-163/Domain: pre-S2 domain #status predicted <PRE2>
F;164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

  Query Match            97.6%;  Score 1883;  DB 2;  Length 389;
  Best Local Similarity  98.8%;  Pred. No. 6.3e-127;
  Matches  338;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||||||| | :||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASHLSSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 3
S32202
large surface antigen - hepatitis B virus (subtype ayw, isolate patient C1005)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, isolate patient C1005
C;Date: 20-Feb-1995 #sequence_revision 06-Dec-1996 #text_change 06-Dec-1996
C;Accession: S32202
R;Preisler-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Rasenack, J.
submitted to the EMBL Data Library, March 1993
A;Description: Identification and sequence analysis of hepatitis B virus DNA in immunolog
A;Reference number: S32202
A;Accession: S32202
A;Molecule type: DNA
A;Residues: 1-389 <PRE>
A;Cross-references: EMBL:X72702
A;Experimental source: subtype ayw, isolate patient C1005
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-389/Product: large surface antigen (gene pre-S1/pre-S2/S) #status predicted <DSL>
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>

  Query Match            97.4%;  Score 1879;  DB 2;  Length 389;
  Best Local Similarity  98.5%;  Pred. No. 1.2e-126;
  Matches  337;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||||::||||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGIIQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTAGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSAGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 TTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 4
S47407
surface antigen - hepatitis B virus (subtype ayw4)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (midc
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw4
C;Date: 23-Nov-1994 #sequence_revision 12-May-1995 #text_change 26-Aug-1999
C;Accession: S47407; JQ2071
R;Pluciennicczak, A.
submitted to the EMBL Data Library, August 1994
A;Description: Molecular cloning and sequencing of two complete genomes of polish isolate
A;Reference number: S47404
A;Accession: S47407
A;Molecule type: DNA
A;Residues: 1-389 <PLU>
A;Cross-references: EMBL:Z35716; NID:g527435; PIDN:CAA84788.1; PID:g527439
A;Experimental source: subtype ayw4
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Contents: genogroup D
A;Accession: JQ2071
A;Molecule type: DNA
A;Residues: 164-389 <NOR>
```

```
A;Experimental source: subtype ayw3, strain 8950/90
C;Genetics:
A;Gene: S
A;Introns: 111/3
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-108/Domain: pre-S1 domain #status predicted <PRE1>
F;109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
P;109-163/Domain: pre-S2 domain #status predicted <PRE2>
F;164-389/Product: surface antigen S (small envelope protein) #status predicted <SAG>

  Query Match             97.2%;  Score 1875;  DB 2;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 2.3e-126;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGIIQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSVGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTTVQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 5
SAVLAH
large surface antigen - hepatitis B virus (subtype ayw)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 07-May-1999
C;Accession: A03703; JQ2064; PQ0591
R;Galibert, F.; Mandart, E.; Fitoussi, F.; Tiollais, P.; Charnay, P.
Nature 281, 646-650, 1979
A;Title: Nucleotide sequence of the hepatitis B virus genome (subtype ayw) in E. coli.
A;Reference number: A93214; MUID:81012091; PMID:399327
A;Accession: A03703
A;Molecule type: DNA
A;Residues: 1-389 <GAL>
A;Cross-references: GB:J02203
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Accession: JQ2064
A;Molecule type: DNA
A;Residues: 164-389 <NOR>
A;Experimental source: genogroup D, subtype ayw3
R;Norder, H.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 3141-3145, 1992
A;Title: Molecular basis of hepatitis B virus serotype variations within the four major
A;Reference number: PQ0453; MUID:93107848; PMID:1469353
A;Accession: PQ0591
A;Molecule type: DNA
A;Residues: 264-343 <NO2>
A;Experimental source: subtype ayw3, Kli
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             96.8%;  Score 1868;  DB 1;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 7.3e-126;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||||||| || ||:||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 6
SAVLA1
large surface antigen - hepatitis B virus (strain alpha1)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C;Accession: B34773; JQ2082
R;Tong, S.; Li, J.; Vitvitski, L.; Trepo, C.
Virology 176, 596-603, 1990
A;Title: Active hepatitis B virus replication in the presence of anti-HBe is associated v
A;Reference number: A34773; MUID:90266476; PMID:2345966
A;Accession: B34773
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-389 <TON>
A;Cross-references: EMBL:M32138; NID:g329667; PIDN:AAA45502.1; PID:g329669
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Accession: JQ2082
A;Molecule type: DNA
A;Residues: 164-389 <NOR>
A;Experimental source: genogroup D, subtype ayw2, strain HBV-alpha
C;Genetics:
A;Gene: pre-S1/pre-S2/S
A;Introns: 111/3
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             96.8%;  Score 1868;  DB 1;  Length 389;
  Best Local Similarity   97.7%;  Pred. No. 7.3e-126;
  Matches  334;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYLPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||:||||||:||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTKILTIPKSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           ||||||||  |||||||||| |::||||||||||||||||||||
Db     348 GLSPTVWLLVIWMMWYWGPKLPTILSPFLPLLPIFFCLWVYI 389


RESULT 7
S20753
surface antigen - hepatitis B virus (subtype ayw, patient E)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient E
C;Date: 20-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
C;Accession: S20753
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submitted to the EMBL Data Library, March 1992
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati
A;Reference number: S20745
A;Accession: S20753
A;Molecule type: DNA
A;Residues: 1-389 <LAI>
A;Cross-references: EMBL:X65259; NID:g59439; PIDN:CAA46357.1; PID:g59440
A;Experimental source: subtype ayw, patient E
C;Genetics:
A;Gene: S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-108/Domain: pre-S1 domain #status predicted <PRE1>
F;109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
F;109-163/Domain: pre-S2 domain #status predicted <PRE2>
F;164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

  Query Match             96.5%;  Score 1861;  DB 2;  Length 389;
  Best Local Similarity   97.7%;  Pred. No. 2.3e-125;
  Matches  334;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           |||||||||||||||||||||||||||| |:||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILHTVPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||||||||||||||||| || ||:||| :| ||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSISARTGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           .||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           ||||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 8
SAVLAJ
large surface antigen - hepatitis B virus (subtype adyw)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Nov-1996
C;Accession: A93217; A03703
R;Pasek, M.; Goto, T.; Gilbert, W.; Zink, B.; Schaller, H.; MacKay, P.; Leadbetter, G.; M
Nature 282, 575-579, 1979
A;Title: Hepatitis B virus genes and their expression in E. coli.
A;Reference number: A93217; MUID:81012115; PMID:399329
A;Accession: A93217
A;Molecule type: DNA
A;Residues: 1-389 <PAS>
A;Cross-references: GB:J02202
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             96.0%;  Score 1851;  DB 1;  Length 389;
  Best Local Similarity   97.1%;  Pred. No. 1.2e-124;
  Matches  332;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||||||||||||||||||||||||||::|||||||||||||||||||||||| ||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGIMQTLPANPPPASTNRQSGRQPTPLSPPLRTTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           || ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     108 AMHWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTTSPISSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPISNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGSCRTC 245
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGSCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTPAQGISMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||| |||||||||||||||||||||||||||||||||| ||
Db     348 GLSPIVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWAYI 389


RESULT 9
S41871
surface antigen - hepatitis B virus (subtype ayw, patient P)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient P
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999
C;Accession: S41871
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.
submitted to the EMBL Data Library, January 1994
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg chroni
```

A;Reference number: S41869
A;Accession: S41871
A;Molecule type: DNA
A;Residues: 1-389 <LAI>
A;Cross-references: EMBL:X77310; NID:g452615; PIDN:CAA54516.1; PID:g452616
A;Experimental source: subtype ayw, patient P
C;Genetics:
A;Gene: S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-108/Domain: pre-S1 domain #status predicted <PRE1>
F;109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
F;109-163/Domain: pre-S2 domain #status predicted <PRE2>
F;164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 93.5%; Score 1803; DB 2; Length 389;
Best Local Similarity 95.3%; Pred. No. 3.1e-121;
Matches 326; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||||||:|:||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTVPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||||||| |||||:| |:|   ||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSHISSIFARTGEPVTIMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPQLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMSRRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVGFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          ||||||||  |||||||| |||||||||||||||||||||||
Db    348 GLSPTVWLLGIWMMWYWGHSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 10
S41869
surface antigen - hepatitis B virus (subtype ayw, patient A) (fragment)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient A
C;Date: 06-Jan-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
C;Accession: S41869
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.
submitted to the EMBL Data Library, January 1994
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg chroni
A;Reference number: S41869
A;Accession: S41869
A;Molecule type: DNA
A;Residues: 1-378 <LAI>
A;Cross-references: EMBL:X77309; NID:g452611; PIDN:CAA54515.1; PID:g452612
A;Experimental source: subtype ayw, patient A
C;Genetics:
A;Gene: S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-378/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-108/Domain: pre-S1 domain #status predicted <PRE1>
F;109-378/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
F;109-163/Domain: pre-S2 domain #status predicted <PRE2>
F;164-378/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 93.3%; Score 1799; DB 2; Length 378;
Best Local Similarity 97.9%; Pred. No. 5.9e-121;
Matches 324; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||||||:|:||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTVPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||||||| || ||:|||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPL 336
          |||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPL 378
```

RESULT 11
S20749
surface antigen - hepatitis B virus (subtype ayw, patient CI)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen S (small en
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient CI
C;Date: 20-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
C;Accession: S20749
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submitted to the EMBL Data Library, March 1992
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati
A;Reference number: S20745
A;Accession: S20749
A;Molecule type: DNA
A;Residues: 1-389 <LAI>
A;Cross-references: EMBL:X65258; NID:g59434; PIDN:CAA46353.1; PID:g59435
A;Experimental source: subtype ayw, patient CI
A;Note: typical methionine for the beginning of pre-S2 domain at position 109 is missing
C;Genetics:
A;Gene: S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 93.2%; Score 1797; DB 2; Length 389;
Best Local Similarity 93.0%; Pred. No. 8.4e-121;
Matches 319; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

```
Qy      5 DGAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHP 64
          ||||||||| ||||||||||||||||||| |:|||||||||||:|||||||||||||:|||
Db     47 DGAGAFGLGLTPPHGGLLGWSPQAQGILHTVPANPPPASTNRQTGRQPTPLSPPLRDTHP 106

Qy     65 QAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMEN 124
          ||:|||||||||||||||||||||||||||||||||||||||||| ||:||||||||  ||||
Db    107 QAVQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPLSSIFSRIGDPVTNMEN 166

Qy    125 ITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPT 184
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    167 ITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFRGGTTVCLGQNSQSPTSNHSPT 226

Qy    185 SCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRT 244
          ||||||||||||||| |||||||||||||||||:||||| |||||||::|||||||:|
```

```
Db    227 SCPPTCPGYRWMCLRGFIIFLFILLLCLIFLLVLLEYQGMLHVCPLIPGTTTTSTGPCKT 286

Qy    245 CTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWF 304
          ||||||| ||:||||||| |||||||||||||||| |:|||||| ||||||||||||||||
Db    287 CTTPAQGNSMFPSCCCTKTSDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWF 346

Qy    305 VGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||| ||||||||||||||||||||||||||||||||||
Db    347 VGLSPTVWLSAIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 12
SAVLA
large surface antigen - hepatitis B virus (subtype adr)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 07-May-1999
C;Accession: A03705; S04569; JQ2107; PQ0608
R;Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Sugino, Y.; Nishioka, K.
Nucleic Acids Res. 11, 1747-1757, 1983
A;Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype
A;Reference number: A93460; MUID:83168919; PMID:6300776
A;Accession: A03705
A;Molecule type: DNA
A;Residues: 1-400 <ONO>
A;Cross-references: GB:V00867
R;Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.
Nucleic Acids Res. 17, 2124, 1989
A;Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtyp
A;Reference number: S04568; MUID:89183619; PMID:2928116
A;Accession: S04569
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-50,'R',52-66,'YP',69-129,'V',131-142,'P',144-164,'S',166-176,'S',178-338,'
A;Cross-references: EMBL:X14193
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Accession: JQ2107
A;Molecule type: DNA
A;Residues: 175-400 <NOR>
A;Experimental source: genogroup C, subtype adr, strain pBRHBadr4
R;Norder, H.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 3141-3145, 1992
A;Title: Molecular basis of hepatitis B virus serotype variations within the four major
A;Reference number: PQ0453; MUID:93107848; PMID:1469353
A;Accession: PQ0608
A;Molecule type: DNA
A;Residues: 275-354 <NO2>
A;Experimental source: subtype adrq+, Bau
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;120-400/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;175-400/Product: major surface antigen (gene S) #status predicted <MSA>
F;15,123,177/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             91.8%;  Score 1771;  DB 1;  Length 400;
  Best Local Similarity   92.1%;  Pred. No. 6.1e-119;
  Matches  315;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||| |||||||||||||||||||| |:|| |||||||||||||||:||||::|||
Db     59 GAGAFGPGFTPPHGGLLGWSPQAQGILTTVPAAPPPASTNRQSGRQPTPISPPLRDSHPQ 118

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||| | |||||||||||||||||||||||||| ||||||||| |||| ||||
Db    119 AMQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRTGDPAPNMENT 178

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||    | ||||||||||||||
Db    179 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 238

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||| |||||||||||||||||||||||||||||||||||||||||||:||:||||||||:||
Db    239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 298

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | |||||||:|||||||||||||||||||||||| :||||||| ||||||||||||||||
Db    299 TIPAQGTSMFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFV 358

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||:||||||||||||||||||||
Db    359 GLSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 400


RESULT 13
SAVLCP
large surface antigen - hepatitis B virus
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jun-2000
C;Accession: C28885; JQ2123
R;Vaudin, M.; Wolstenholme, A.J.; Tsiquaye, K.N.; Zuckerman, A.J.; Harrison, T.J.
J. Gen. Virol. 69, 1383-1389, 1988
A;Title: The complete nucleotide sequence of the genome of a hepatitis B virus isolated f
A;Reference number: A92796; MUID:88258473; PMID:2838576
A;Accession: C28885
A;Molecule type: DNA
A;Residues: 1-389 <VAU>
A;Cross-references: GB:D00220; NID:g221505; PIDN:BAA00159.1; PID:g221508
A;Experimental source: strain LSH, chimpanzee
A;Note: the authors translated the codon GAG for residue 327 as Gln
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Accession: JQ2123
A;Molecule type: DNA
A;Residues: 164-389 <NOR>
A;Experimental source: subtype adw2, strain adwLSH
C;Genetics:
A;Gene: pre-S1/pre-S2/S; S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             91.0%;  Score 1755;  DB 1;  Length 389;
  Best Local Similarity   91.5%;  Pred. No. 8.2e-118;
  Matches  313;  Conservative    9;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||| |||||||||| |||||||| |||||||||||||||||||||||||:||||
Db     48 GAGAFGPGFTPPHGGLLGLSPQAQGILTTLPANPPPASTNRQSGRQPTPLSPPLRDTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||| |||||||||||||||||||:|||| | | |||:||  |||| |||||
Db    108 AMQWNSTTFHQALQDPRVRGLYFPAGGSSSGTLNPVPNTASHISSVFSTTGDPAPNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||:|||||||||||||||||  ||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTKILTIPQSLDSWWTSLNFLGGAPVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db    228 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCKTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||||||||: |||||||||||||||||||||||||||| |||||||| ||||||| ||||
Db    288 TTPAQGTSLIPSCCCTKPSDGNCTCIPIPSSWAFAKFLWEWASVRFSWLSLLAPFVQWFA 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
```

```
          |||||||| |||||||||:||:|||||:|||||||||||||||
Db    348 GLSPTVWLLAIWMMWYWGPNLYNILSPFIPLLPIFFCLWVYI 389
```

RESULT 14
S35528
surface antigen - hepatitis B virus (subtype adr)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype adr
C;Date: 09-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C;Accession: S35528
R;Mukaide, M.; Kumazawa, T.; Hoshi, A.; Kawaguchi, R.; Hikiji, K.
Nucleic Acids Res. 20, 6105, 1992
A;Title: The complete nucleotide sequence of hepatitis B virus, subtype adr (SRADR) and
A;Reference number: S35527; MUID:93096607; PMID:1461746
A;Accession: S35528
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-400 <MUK>
A;Cross-references: EMBL:D12980; NID:g221500; PIDN:BAA02359.1; PID:g221502
A;Experimental source: subtype adr
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992
C;Genetics:
A;Gene: S
A;Introns: 165/2
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-400/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-119/Domain: pre-S1 domain #status predicted <PRE1>
F;120-400/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
F;120-174/Domain: pre-S2 domain #status predicted <PRE2>
F;175-400/Product: surface antigen S (small envelope protein) #status predicted <PSD>

  Query Match 90.5%; Score 1745; DB 2; Length 400;
  Best Local Similarity 90.9%; Pred. No. 4.3e-117;
  Matches 311; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||| || |||||||||||||||||||| |:|| ||||||||||||||||:|||||::|||
Db     59 GAGDFGPGFTPPHGGLLGWSPQAQGILTTVPAAPPPASTNRQSGRQPTPISPPLRDSHPQ 118

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||| | |||||||||||||||||||||||||| ||||||||| |||| |||:
Db    119 AMQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRTGDPAPNMEST 178

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||    | ||||||||||||||
Db    179 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 238

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||| |||||||||||||||||||||||||||||||||||||||||:||:|||||||:||
Db    239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 298

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | |||||||:||| |||||||||||||||||||| :||||||| |||||||||||||||
Db    299 TIPAQGTSMFPSCWCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFA 358

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||:||||||||||||||||||||
Db    359 GLSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 400
```

RESULT 15
S41870
surface antigen - hepatitis B virus (subtype ayw, patient M) (fragment)
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient M
C;Date: 06-Jan-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
C;Accession: S41870
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.
submitted to the EMBL Data Library, January 1994
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg chroni
A;Reference number: S41869
A;Accession: S41870
A;Molecule type: DNA
A;Residues: 1-378 <LAI>
A;Cross-references: EMBL:X77308; NID:g452613; PIDN:CAA54514.1; PID:g452614
A;Experimental source: subtype ayw, patient M
C;Genetics:
A;Gene: S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: surface antigen
F;1-378/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F;1-108/Domain: pre-S1 domain #status predicted <PRE1>
F;109-378/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PS
F;109-163/Domain: pre-S2 domain #status predicted <PRE2>
F;164-378/Product: surface antigen S (small envelope protein) #status predicted <PSD>

  Query Match 89.9%; Score 1734; DB 2; Length 378;
  Best Local Similarity 94.0%; Pred. No. 2.5e-116;
  Matches 311; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||||||:|:|||||||||||||||||||||||| ||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTVPANPPPASTNRQSGRQPTPLSPPLRKTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||| |||||||||| || ||:||| :| |||   ||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAFGSSSGTVNPVLTTASPLSSISARTGDPVTIMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||:
Db    288 TTPAQGTSMYPSCCCTKPSHGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFM 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPL 336
          ||||||||| |||||||||||||: ||:||
Db    348 GLSPTVWLSAIWMMWYWGPSLYSIVMPFIPL 378
```

Search completed: February 3, 2003, 09:26:27
Job time : 23.9721 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:24:27 ; Search time 28.4129 Seconds
(without alignments)
988.252 Million cell updates/sec

Title: US-09-890-752A-1
Perfect score: 1929
Sequence: 1 MGRGDGAGAFGLGFTPPHGG.....SILSPFLPLLPIFFCLWVYI 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 404799 seqs, 80919614 residues

Total number of hits satisfying chosen parameters: 404799

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1756	91.0	400	1	PCT-US02-32263-46	Sequence 46, Appl
2	1756	91.0	400	6	US-10-287-994-46	Sequence 46, Appl
3	300	15.6	67	6	US-10-326-908-21	Sequence 21, Appl
4	298	15.4	64	6	US-10-326-908-19	Sequence 19, Appl
5	131	6.8	63	6	US-10-326-908-18	Sequence 18, Appl
6	124	6.4	68	6	US-10-326-908-20	Sequence 20, Appl
7	107	5.5	20	5	US-09-863-054-22	Sequence 22, Appl
8	105	5.4	276	5	US-09-724-676-50955	Sequence 50955, A
9	105	5.4	276	5	US-09-724-676-50956	Sequence 50956, A
10	105	5.4	276	5	US-09-724-676-50957	Sequence 50957, A
11	105	5.4	276	5	US-09-724-676A-50955	Sequence 50955, A
12	105	5.4	276	5	US-09-724-676A-50956	Sequence 50956, A
13	105	5.4	276	5	US-09-724-676A-50957	Sequence 50957, A
14	105	5.4	328	5	US-09-724-676-50953	Sequence 50953, A
15	105	5.4	328	5	US-09-724-676-50954	Sequence 50954, A
16	105	5.4	328	5	US-09-724-676A-50953	Sequence 50953, A
17	105	5.4	328	5	US-09-724-676A-50954	Sequence 50954, A
18	103	5.3	19	5	US-09-863-054-28	Sequence 28, Appl
19	102	5.3	682	6	US-10-144-779-327	Sequence 327, App
20	101.5	5.3	2509	6	US-10-237-271-1	Sequence 1, Appli
21	100.5	5.2	641	6	US-10-311-034-15	Sequence 15, Appl
22	99.5	5.2	845	5	US-09-724-676-94342	Sequence 94342, A
23	99.5	5.2	845	5	US-09-724-676A-94342	Sequence 94342, A
24	99.5	5.2	1078	6	US-10-170-385-295	Sequence 295, App
25	99.5	5.2	1088	6	US-10-284-130-13	Sequence 13, Appl
26	99	5.1	18	5	US-09-863-054-29	Sequence 29, Appl
27	97	5.0	1251	1	PCT-US02-10780-80	Sequence 80, Appl
28	97	5.0	1251	6	US-10-114-270-80	Sequence 80, Appl
29	96	5.0	611	6	US-10-281-346-4	Sequence 4, Appli
30	95.5	5.0	1367	1	PCT-US02-40225-3355	Sequence 3355, Ap
31	95.5	5.0	1367	6	US-10-320-797-3355	Sequence 3355, Ap
32	95.5	5.0	1441	6	US-10-170-682-3	Sequence 3, Appli
33	95	4.9	17	5	US-09-863-054-30	Sequence 30, Appl
34	95	4.9	606	5	US-09-724-676-86512	Sequence 86512, A
35	95	4.9	606	5	US-09-724-676-86513	Sequence 86513, A
36	95	4.9	606	5	US-09-724-676-86514	Sequence 86514, A
37	95	4.9	606	5	US-09-724-676-86515	Sequence 86515, A
38	95	4.9	606	5	US-09-724-676A-86512	Sequence 86512, A
39	95	4.9	606	5	US-09-724-676A-86513	Sequence 86513, A
40	95	4.9	606	5	US-09-724-676A-86514	Sequence 86514, A
41	95	4.9	606	5	US-09-724-676A-86515	Sequence 86515, A
42	95	4.9	798	5	US-09-724-676-86525	Sequence 86525, A
43	95	4.9	798	5	US-09-724-676-86526	Sequence 86526, A
44	95	4.9	798	5	US-09-724-676-86527	Sequence 86527, A
45	95	4.9	798	5	US-09-724-676-86528	Sequence 86528, A

ALIGNMENTS

```
RESULT 1
PCT-US02-32263-46
; Sequence 46, Application PC/TUS0232263
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Bowe, Caryn
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
;  FILE REFERENCE: 040853-01-5050WO
;  CURRENT APPLICATION NUMBER: PCT/US02/32263
;  CURRENT FILING DATE:  2002-10-31
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/334,233
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: US 60/334,301
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
PCT-US02-32263-46

  Query Match             91.0%;  Score 1756;  DB 1;  Length 400;
  Best Local Similarity   91.2%;  Pred. No. 4.5e-106;
  Matches  312;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||| || |||||||||||||||||||||| |:|| ||| |||||||||||||:||||||::|||
Db      59 GAGDFGPGFTPPHGGLLGWSPQAQGILTTVPAAPPPVSTNRQSGRQPTPISPPLRDSHPQ 118
```

```
Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||| | |||||||||||||||||||||||||||||||||||| |||| |||:
Db     119 AMQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRTGDPAPNMEST 178

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||   | ||||||||||||||
Db     179 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 238

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| |||||||||||||||||||||||||||||||||||||||||:||:|||||||||:||
Db     239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 298

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||:||||||||||||||||||||| :||||||| ||||||||||||||||
Db     299 TIPAQGTSMFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFA 358

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||:||||||||||||||||||||
Db     359 GLSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 400


RESULT 2
US-10-287-994-46
; Sequence 46, Application US/10287994
; GENERAL INFORMATION:
;  APPLICANT: Neose Technologies, Inc.
;  APPLICANT:  DeFrees, Shawn
;  APPLICANT:  Zopf, David
;  APPLICANT:  Bayer, Robert
;  APPLICANT:  Bowe, Caryn
;  APPLICANT:  Hakes, David
;  APPLICANT:  Chen, Xi
;  TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
;  FILE REFERENCE: 040853-01-5052-00
;  CURRENT APPLICATION NUMBER: US/10/287,994
;  CURRENT FILING DATE:  2002-11-05
;  PRIOR APPLICATION NUMBER: US 60/328,523
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/344,692
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/387,292
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: US 60/391,777
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/396,594
;  PRIOR FILING DATE: 2002-07-17
;  PRIOR APPLICATION NUMBER: US 60/404,249
;  PRIOR FILING DATE: 2002-08-16
;  PRIOR APPLICATION NUMBER: US 60/407,527
;  PRIOR FILING DATE: 2002-08-28
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-287-994-46

  Query Match             91.0%;  Score 1756;  DB 6;  Length 400;
  Best Local Similarity   91.2%;  Pred. No. 4.5e-106;
  Matches  312;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||| || |||||||||||||||||||| |:|| ||| |||||||||||:|||||::|||
Db      59 GAGDFGPGFTPPHGGLLGWSPQAQGILTTVPAAPPPVSTNRQSGRQPTPISPPLRDSHPQ 118

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||| | |||||||||||||||||||||||||||||||||||| |||| |||:
Db     119 AMQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRTGDPAPNMEST 178

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||   | ||||||||||||||
Db     179 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 238

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| |||||||||||||||||||||||||||||||||||||||||:||:|||||||||:||
Db     239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 298

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||:||||||||||||||||||||| :||||||| ||||||||||||||||
Db     299 TIPAQGTSMFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFA 358

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||:||||||||||||||||||||
Db     359 GLSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 400


RESULT 3
US-10-326-908-21
; Sequence 21, Application US/10326908
; GENERAL INFORMATION:
;  APPLICANT: Christopher JONES
;  APPLICANT:  Andrew BACON
;  APPLICANT:  Gill DOUCE
;  APPLICANT:  Mark PAGE
;  TITLE OF INVENTION: Designing Immunogens
;  FILE REFERENCE: 117-433 / N78027D
;  CURRENT APPLICATION NUMBER: US/10/326,908
;  CURRENT FILING DATE:  2002-12-23
;  PRIOR APPLICATION NUMBER: US 10/110,036
;  PRIOR FILING DATE: 2002-04-08
;  PRIOR APPLICATION NUMBER: PCT/GB00/03857
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: GB 0007789.1
;  PRIOR FILING DATE: 2000-03-30
;  PRIOR APPLICATION NUMBER: GB 9923902.2
;  PRIOR FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: MS Word
; SEQ ID NO 21
;   LENGTH: 67
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide immunogen
US-10-326-908-21

  Query Match             15.6%;  Score 300;  DB 6;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 8.3e-13;
  Matches   57;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      65 QAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALN 121
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6 QAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALN 62


RESULT 4
US-10-326-908-19
; Sequence 19, Application US/10326908
; GENERAL INFORMATION:
;  APPLICANT: Christopher JONES
;  APPLICANT:  Andrew BACON
;  APPLICANT:  Gill DOUCE
;  APPLICANT:  Mark PAGE
;  TITLE OF INVENTION: Designing Immunogens
;  FILE REFERENCE: 117-433 / N78027D
;  CURRENT APPLICATION NUMBER: US/10/326,908
;  CURRENT FILING DATE:  2002-12-23
;  PRIOR APPLICATION NUMBER: US 10/110,036
;  PRIOR FILING DATE: 2002-04-08
;  PRIOR APPLICATION NUMBER: PCT/GB00/03857
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: GB 0007789.1
;  PRIOR FILING DATE: 2000-03-30
;  PRIOR APPLICATION NUMBER: GB 9923902.2
;  PRIOR FILING DATE: 1999-10-08
```

```
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: MS Word
; SEQ ID NO 19
;   LENGTH: 64
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide immunogen
US-10-326-908-19

  Query Match             15.4%;  Score 298;  DB 6;  Length 64;
  Best Local Similarity   96.6%;  Pred. No. 1.1e-12;
  Matches   57;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      65 QAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNME 123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db       6 QAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALKSE 64


RESULT 5
US-10-326-908-18
; Sequence 18, Application US/10326908
; GENERAL INFORMATION:
;  APPLICANT: Christopher JONES
;  APPLICANT:  Andrew BACON
;  APPLICANT:  Gill DOUCE
;  APPLICANT:  Mark PAGE
;  TITLE OF INVENTION: Designing Immunogens
;  FILE REFERENCE: 117-433 / N78027D
;  CURRENT APPLICATION NUMBER: US/10/326,908
;  CURRENT FILING DATE:  2002-12-23
;  PRIOR APPLICATION NUMBER: US 10/110,036
;  PRIOR FILING DATE: 2002-04-08
;  PRIOR APPLICATION NUMBER: PCT/GB00/03857
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: GB 0007789.1
;  PRIOR FILING DATE: 2000-03-30
;  PRIOR APPLICATION NUMBER: GB 9923902.2
;  PRIOR FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: MS Word
; SEQ ID NO 18
;   LENGTH: 63
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide immunogen
US-10-326-908-18

  Query Match             6.8%;  Score 131;  DB 6;  Length 63;
  Best Local Similarity   91.7%;  Pred. No. 0.064;
  Matches   22;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQ 29
           |||||||||||||||||||||||::
Db      40 GAGAFGLGFTPPHGGLLGWSPQSE 63


RESULT 6
US-10-326-908-20
; Sequence 20, Application US/10326908
; GENERAL INFORMATION:
;  APPLICANT: Christopher JONES
;  APPLICANT:  Andrew BACON
;  APPLICANT:  Gill DOUCE
;  APPLICANT:  Mark PAGE
;  TITLE OF INVENTION: Designing Immunogens
;  FILE REFERENCE: 117-433 / N78027D
;  CURRENT APPLICATION NUMBER: US/10/326,908
;  CURRENT FILING DATE:  2002-12-23
;  PRIOR APPLICATION NUMBER: US 10/110,036
;  PRIOR FILING DATE: 2002-04-08
;  PRIOR APPLICATION NUMBER: PCT/GB00/03857
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: GB 0007789.1
;  PRIOR FILING DATE: 2000-03-30
;  PRIOR APPLICATION NUMBER: GB 9923902.2
;  PRIOR FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: MS Word
; SEQ ID NO 20
;   LENGTH: 68
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide immunogen
US-10-326-908-20

  Query Match             6.4%;  Score 124;  DB 6;  Length 68;
  Best Local Similarity   88.0%;  Pred. No. 0.19;
  Matches   22;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQG 30
           ||||||||||||||||||||||   |
Db      40 GAGAPGLGFTPPHGGLLGWSPHDIG 64


RESULT 7
US-09-863-054-22
; Sequence 22, Application US/09863054
;  GENERAL INFORMATION:
;       APPLICANT: Chisari, Francis V.
;       TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
;                           Lymphocyte Responses to Hepatitis B Virus
;       NUMBER OF SEQUENCES: 57
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Townsend and Townsend and Crew LLP
;            STREET: Two Embarcadero Center, Eighth Floor
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94111-3834
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/863,054
;            FILING DATE: 21-May-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 07/935,898
;            FILING DATE: 26-AUG-1992
;            APPLICATION NUMBER: US 08/024,120
;            FILING DATE: 26-FEB-1993
;            APPLICATION NUMBER: US 08/396,283
;            FILING DATE: 27-FEB-1995
;            APPLICATION NUMBER: US 08/463,486
;            FILING DATE: 05-JUN-1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Lockyer, Jean M.
;            REGISTRATION NUMBER: 44,879
;            REFERENCE/DOCKET NUMBER: 014740-000421US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 576-0200
;            TELEFAX: (415) 576-0300
;   INFORMATION FOR SEQ ID NO: 22:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 20 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-863-054-22
```

```
  Query Match             5.5%;  Score 107;  DB 5;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.8;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     287 ASARFSWLSLLVPFVQWFVG 306
           ||||||||||||||||||||
Db       1 ASARFSWLSLLVPFVQWFVG 20
```

RESULT 8
US-09-724-676-50955
; Sequence 50955, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50955
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-50955

```
  Query Match             5.4%;  Score 105;  DB 5;  Length 276;
  Best Local Similarity   26.0%;  Pred. No. 12;
  Matches   57;  Conservative   11;  Mismatches   73;  Indels   78;  Gaps   10;

Qy       2 GRGDGAGAFGLGFTPPHGGLL-----------------GWSPQAQGILETLPANPP--- 40
           | | | | :|| |  |  |||                 |  |         |  ||
Db      32 GGGGGGGSFGGGGPPGLGGLFQAGMPKLRSTANRDNDSGGSRP---------PLLPPGGR 82

Qy      41 -----PASTNRQSGRQPTPLSPPLRNTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSS 95
                 | |     || | | ||  |:  |:  :         :  ||       |  ||
Db      83 STSAKPFSPPSGPGRFPVP-SPGHRSGPPEPQR-------NRMPPPR------PDVGSKP 128

Qy      96 GTV-NPVPTTVSPISSIFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLD 154
             ::  |||:|  || |     | |      : |                        |:
Db     129 DSIPPPVPSTPRPIQSSLHNRGSPP-----VPGG--------------------PRQPS 162

Qy     155 SWWTSLNFLGGTTVCLGQNS--QSPTSNHSPTSCPPTCP 191
             |   | |     ||  |  ||| |: || |  |  |
Db     163 PGPTPPPPPGNRGTALGGGSIRQSPLSSSSPFSNRPPLP 201
```

RESULT 9
US-09-724-676-50956
; Sequence 50956, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50956
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-50956

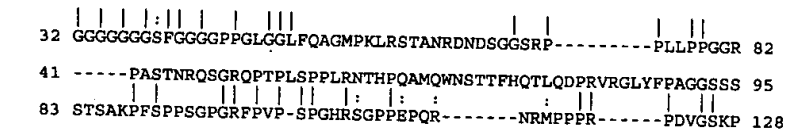
```
  Query Match             5.4%;  Score 105;  DB 5;  Length 276;
  Best Local Similarity   26.0%;  Pred. No. 12;
  Matches   57;  Conservative   11;  Mismatches   73;  Indels   78;  Gaps   10;

Qy       2 GRGDGAGAFGLGFTPPHGGLL-----------------GWSPQAQGILETLPANPP--- 40
           | | | | :|| |  |  |||                 |  |         |  ||
Db      32 GGGGGGGSFGGGGPPGLGGLFQAGMPKLRSTANRDNDSGGSRP---------PLLPPGGR 82

Qy      41 -----PASTNRQSGRQPTPLSPPLRNTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSS 95
                 | |     || | | ||  |:  |:  :         :  ||       |  ||
Db      83 STSAKPFSPPSGPGRFPVP-SPGHRSGPPEPQR-------NRMPPPR------PDVGSKP 128

Qy      96 GTV-NPVPTTVSPISSIFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLD 154
             ::  |||:|  || |     | |      : |                        |:
Db     129 DSIPPPVPSTPRPIQSSLHNRGSPP-----VPGG--------------------PRQPS 162

Qy     155 SWWTSLNFLGGTTVCLGQNS--QSPTSNHSPTSCPPTCP 191
             |   | |     ||  |  ||| |: || |  |  |
Db     163 PGPTPPPPPGNRGTALGGGSIRQSPLSSSSPFSNRPPLP 201
```

RESULT 10
US-09-724-676-50957
; Sequence 50957, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50957
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-50957

```
  Query Match             5.4%;  Score 105;  DB 5;  Length 276;
  Best Local Similarity   26.0%;  Pred. No. 12;
  Matches   57;  Conservative   11;  Mismatches   73;  Indels   78;  Gaps   10;

Qy       2 GRGDGAGAFGLGFTPPHGGLL-----------------GWSPQAQGILETLPANPP--- 40
           | | | | :|| |  |  |||                 |  |         |  ||
Db      32 GGGGGGGSFGGGGPPGLGGLFQAGMPKLRSTANRDNDSGGSRP---------PLLPPGGR 82

Qy      41 -----PASTNRQSGRQPTPLSPPLRNTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSS 95
                 | |     || | | ||  |:  |:  :         :  ||       |  ||
Db      83 STSAKPFSPPSGPGRFPVP-SPGHRSGPPEPQR-------NRMPPPR------PDVGSKP 128

Qy      96 GTV-NPVPTTVSPISSIFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLD 154
             ::  |||:|  || |     | |      : |                        |:
Db     129 DSIPPPVPSTPRPIQSSLHNRGSPP-----VPGG--------------------PRQPS 162

Qy     155 SWWTSLNFLGGTTVCLGQNS--QSPTSNHSPTSCPPTCP 191
             |   | |     ||  |  ||| |: || |  |  |
Db     163 PGPTPPPPPGNRGTALGGGSIRQSPLSSSSPFSNRPPLP 201
```

RESULT 11
US-09-724-676A-50955
; Sequence 50955, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50955
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-50955

```
  Query Match             5.4%;  Score 105;  DB 5;  Length 276;
  Best Local Similarity   26.0%;  Pred. No. 12;
  Matches   57;  Conservative   11;  Mismatches   73;  Indels   78;  Gaps   10;

Qy       2 GRGDGAGAFGLGFTPPHGGLL-----------------GWSPQAQGILETLPANPP--- 40
```



```
Db      32 GGGGGGGSFGGGGPPGLGGLFQAGMPKLRSTANRDNDSGGSRP---------PLLPPGGR 82

Qy      41 -----PASTNRQSGRQPTPLSPPLRNTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSS 95
Db      83 STSAKPFSPPSGPGRFPVP-SPGHRSGPPEPQR-------NRMPPPR------PDVGSKP 128

Qy      96 GTV-NPVPTTVSPISSIFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLD 154
Db     129 DSIPPPVPSTPRPIQSSLHNRGSPP-----VPGG--------------------PRQPS 162

Qy     155 SWWTSLNFLGGTTVCLGQNS--QSPTSNHSPTSCPPTCP 191
Db     163 PGPTPPPFPGNRGTALGGGSIRQSPLSSSSPFSNRPPLP 201
```

RESULT 12
US-09-724-676A-50956
; Sequence 50956, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50956
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-50956

Query Match 5.4%; Score 105; DB 5; Length 276;
Best Local Similarity 26.0%; Pred. No. 12;
Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

```
Qy       2 GRGDGAGAFGLGFTPPHGGLL-----------------GWSPQAQGILETLPANPP--- 40
Db      32 GGGGGGGSFGGGGPPGLGGLFQAGMPKLRSTANRDNDSGGSRP---------PLLPPGGR 82

Qy      41 -----PASTNRQSGRQPTPLSPPLRNTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSS 95
Db      83 STSAKPFSPPSGPGRFPVP-SPGHRSGPPEPQR-------NRMPPPR------PDVGSKP 128

Qy      96 GTV-NPVPTTVSPISSIFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLD 154
Db     129 DSIPPPVPSTPRPIQSSLHNRGSPP-----VPGG--------------------PRQPS 162

Qy     155 SWWTSLNFLGGTTVCLGQNS--QSPTSNHSPTSCPPTCP 191
Db     163 PGPTPPPFPGNRGTALGGGSIRQSPLSSSSPFSNRPPLP 201
```

RESULT 13
US-09-724-676A-50957
; Sequence 50957, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50957
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-50957

Query Match 5.4%; Score 105; DB 5; Length 276;
Best Local Similarity 26.0%; Pred. No. 12;
Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

```
Qy       2 GRGDGAGAFGLGFTPPHGGLL-----------------GWSPQAQGILETLPANPP--- 40
Db      32 GGGGGGGSFGGGGPPGLGGLFQAGMPKLRSTANRDNDSGGSRP---------PLLPPGGR 82

Qy      41 -----PASTNRQSGRQPTPLSPPLRNTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSS 95
Db      83 STSAKPFSPPSGPGRFPVP-SPGHRSGPPEPQR-------NRMPPPR------PDVGSKP 128

Qy      96 GTV-NPVPTTVSPISSIFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLD 154
Db     129 DSIPPPVPSTPRPIQSSLHNRGSPP-----VPGG--------------------PRQPS 162

Qy     155 SWWTSLNFLGGTTVCLGQNS--QSPTSNHSPTSCPPTCP 191
Db     163 PGPTPPPFPGNRGTALGGGSIRQSPLSSSSPFSNRPPLP 201
```

RESULT 14
US-09-724-676-50953
; Sequence 50953, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50953
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-50953

Query Match 5.4%; Score 105; DB 5; Length 328;
Best Local Similarity 26.0%; Pred. No. 14;
Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

```
Qy       2 GRGDGAGAFGLGFTPPHGGLL-----------------GWSPQAQGILETLPANPP--- 40
Db      84 GGGGGGGSFGGGGPPGLGGLFQAGMPKLRSTANRDNDSGGSRP---------PLLPPGGR 134

Qy      41 -----PASTNRQSGRQPTPLSPPLRNTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSS 95
Db     135 STSAKPFSPPSGPGRFPVP-SPGHRSGPPEPQR-------NRMPPPR------PDVGSKP 180

Qy      96 GTV-NPVPTTVSPISSIFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLD 154
Db     181 DSIPPPVPSTPRPIQSSLHNRGSPP-----VPGG--------------------PRQPS 214

Qy     155 SWWTSLNFLGGTTVCLGQNS--QSPTSNHSPTSCPPTCP 191
Db     215 PGPTPPPFPGNRGTALGGGSIRQSPLSSSSPFSNRPPLP 253
```

RESULT 15
US-09-724-676-50954
; Sequence 50954, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50954
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-50954

Query Match 5.4%; Score 105; DB 5; Length 328;
Best Local Similarity 26.0%; Pred. No. 14;
Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

```
Qy       2 GRGDGAGAPGLGFTPPHGGLL-----------------GWSPQAQGILETLPANPP--- 40
Db      84 GGGGGGGSFGGGGPPGLGGLFQAGMPKLRSTANRDNDSGGSRP---------PLLPPGGR 134

Qy      41 -----PASTNRQSGRQPTPLSPPLRNTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSS 95
Db     135 STSAKPFSPPSGPGRFPVP-SPGHRSGPPEPQR-------NRMPPPR------PDVGSKP 180

Qy      96 GTV-NPVPTTVSPISSIFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLD 154
Db     181 DSIPPPVPSTPRPIQSSLHNRGSPP-----VPGG--------------------PRQPS 214

Qy     155 SWWTSLNFLGGTTVCLGQNS--QSPTSNHSPTSCPPTCP 191
Db     215 PGPTPPPFPGNRGTALGGGSIRQSPLSSSSPFSNRPPLP 253
```

Search completed: February 3, 2003, 09:35:00
Job time : 30.4129 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:23:57 ; Search time 247.857 Seconds
(without alignments)
902.626 Million cell updates/sec

Title: US-09-890-752A-1
Perfect score: 1929
Sequence: 1 MGRGDGAGAFGLGFTPPHGG.....SILSPFLPLLPIFFCLWVYI 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1929	100.0	347	22	US-09-890-752A-1	Sequence 1, Appli
2	1879	97.4	389	21	US-09-718-095-18	Sequence 18, Appl
3	1870	96.9	389	3	US-07-847-951A-216	Sequence 216, App
4	1870	96.9	389	3	US-07-847-951A-219	Sequence 219, App
5	1870	96.9	389	4	US-08-036-218-214	Sequence 214, App
6	1870	96.9	389	4	US-08-036-218-217	Sequence 217, App
7	1868	96.8	389	7	US-08-360-107-116	Sequence 116, App
8	1868	96.8	389	8	US-08-470-896-106	Sequence 106, App
9	1868	96.8	389	8	US-08-471-913-106	Sequence 106, App
10	1868	96.8	389	8	US-08-475-668-106	Sequence 106, App
11	1868	96.8	389	8	US-08-484-223-106	Sequence 106, App
12	1868	96.8	389	8	US-08-484-223A-106	Sequence 106, App
13	1868	96.8	389	8	US-08-485-546-106	Sequence 106, App
14	1868	96.8	389	8	US-08-485-546A-106	Sequence 106, App
15	1868	96.8	389	8	US-08-485-551-106	Sequence 106, App
16	1868	96.8	389	8	US-08-487-266-106	Sequence 106, App
17	1868	96.8	389	8	US-08-487-266A-106	Sequence 106, App
18	1868	96.8	389	8	US-08-487-355-106	Sequence 106, App
19	1868	96.8	389	8	US-08-487-355A-106	Sequence 106, App
20	1868	96.8	389	13	US-08-919-600-106	Sequence 106, App
21	1868	96.8	389	19	US-09-502-445-106	Sequence 106, App
22	1868	96.8	389	25	US-10-169-668-6	Sequence 6, Appli
23	1868	96.8	389	26	US-10-267-682-106	Sequence 106, App
24	1868	96.8	389	26	US-10-267-748-106	Sequence 106, App
25	1866	96.7	389	1	PCT-US02-09227-2	Sequence 2, Appli
26	1866	96.7	389	22	US-09-821-877-2	Sequence 2, Appli
27	1783	92.4	399	21	US-09-718-095-19	Sequence 19, Appl
28	1745	90.5	400	21	US-09-718-095-17	Sequence 17, Appl
29	1743	90.4	399	21	US-09-718-095-5	Sequence 5, Appli
30	1721	89.2	400	21	US-09-719-528-3	Sequence 3, Appli
31	1721	89.2	400	21	US-09-719-528A-3	Sequence 3, Appli
32	1721	89.2	400	26	US-10-209-264-3	Sequence 3, Appli
33	1700	88.1	400	21	US-09-718-095-15	Sequence 15, Appl
34	1684	87.3	400	21	US-09-718-095-16	Sequence 16, Appl
35	1587	82.3	400	21	US-09-718-095-20	Sequence 20, Appl
36	1539	79.8	281	3	US-07-847-951A-214	Sequence 214, App
37	1539	79.8	281	4	US-08-036-218-212	Sequence 212, App
38	1534	79.5	281	14	US-09-021-769-6	Sequence 6, Appli
39	1534	79.5	281	16	US-09-247-890-12	Sequence 12, Appl
40	1534	79.5	281	21	US-09-724-648-12	Sequence 12, Appl
41	1534	79.5	281	21	US-09-724-890-6	Sequence 6, Appli
42	1534	79.5	281	21	US-09-724-969-12	Sequence 12, Appl
43	1460	75.7	281	9	US-08-537-814A-35	Sequence 35, Appl
44	1460	75.7	281	13	US-08-942-483-35	Sequence 35, Appl
45	1460	75.7	281	15	US-09-182-186A-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-09-890-752A-1
; Sequence 1, Application US/09890752A
; GENERAL INFORMATION:
; APPLICANT: Hildt, Eberhard
; APPLICANT: Hofschneider, Peter
; TITLE OF INVENTION: Particles for Gene Therapy
; FILE REFERENCE: 107070-120 (VOS-013)
; CURRENT APPLICATION NUMBER: US/09/890,752A
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: PCT/DE00/00363
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: DE 199 04 800.2
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein comprising a LHBs and heterologous binding
; OTHER INFORMATION: site RGD
US-09-890-752A-1

Query Match 100.0%; Score 1929; DB 22; Length 347;
Best Local Similarity 100.0%; Pred. No. 4.3e-141;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MGRGDGAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLR 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MGRGDGAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLR 60

Qy      61 NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPAL 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPAL 120

Qy     121 NMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSN 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSN 180

Qy     181 HSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 HSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTG 240

Qy     241 PCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 PCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPF 300

Qy     301 VQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
```

RESULT 2
US-09-718-095-18
; Sequence 18, Application US/09718095
; GENERAL INFORMATION:
; APPLICANT: STUYVER Lieven
; APPLICANT: VAN GEYT Caroline
; APPLICANT: DE GENDT Sija
; TITLE OF INVENTION: New HBV Sequences
; FILE REFERENCE: 2551-52
; CURRENT APPLICATION NUMBER: US/09/718,095
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: EP99870252.6
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US60/169,287
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-718-095-18

Query Match 97.4%; Score 1879; DB 21; Length 389;
Best Local Similarity 98.5%; Pred. No. 3.7e-137;
Matches 337; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||||||||||||||||||||||||||||::|||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGIIQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSAGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 3
US-07-847-951A-216
; Sequence 216, Application US/07847951A
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/847,951A
; FILING DATE: 19920326
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 216:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-847-951A-216

Query Match 96.9%; Score 1870; DB 3; Length 389;
Best Local Similarity 98.0%; Pred. No. 1.8e-136;
Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||||||||||||||||||||||||||||:|||||||||||||||:||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||||||||||||||||||||| |||||||  ||:||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

```
RESULT 4
US-07-847-951A-219
; Sequence 219, Application US/07847951A
;  GENERAL INFORMATION:
;     APPLICANT:  Paoletti, Enzo
;     TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;     TITLE OF INVENTION:  STRAIN
;     NUMBER OF SEQUENCES:  462
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Curtis, Morris & Safford
;       ADDRESSEE:  c/o William S. Frommer
;       STREET:  530 Fifth Avenue
;       CITY:  New York
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/847,951A
;       FILING DATE:  19920326
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Frommer, William S.
;       REGISTRATION NUMBER:  25,506
;       REFERENCE/DOCKET NUMBER:  454310-2400
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 840-3333
;       TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  219:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  389 amino acids
;       TYPE:  AMINO ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-07-847-951A-219

  Query Match             96.9%;  Score 1870;  DB 3;  Length 389;
  Best Local Similarity   98.0%;  Pred. No. 1.8e-136;
  Matches  335;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||||||:|||||||||||||:||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||| ||||||| ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 5
US-08-036-218-214
; Sequence 214, Application US/08036218
;  GENERAL INFORMATION:
;     APPLICANT:  Paoletti, Enzo
;     APPLICANT:  Perkus, Marion E.
;     APPLICANT:  Taylor, Jill
;     APPLICANT:  Tartaglia, James
;     APPLICANT:  Norton, Elizabeth K.
;     APPLICANT:  Riviere, Michel
;     APPLICANT:  de Taisne, Charles
;     APPLICANT:  Limbach, Keith J.
;     APPLICANT:  Johnson, Gerard P.
;     APPLICANT:  Pincus, Steven E.
;     APPLICANT:  Cox, William I.
;     TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE STRAIN
;     NUMBER OF SEQUENCES:  221
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  William S. Frommer c/o
;       ADDRESSEE:  Curtis, Morris & Safford, P.C.
;       STREET:  530 Fifth Avenue
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  United State of America
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/036,218
;       FILING DATE:
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/713,967
;       FILING DATE:  11-JUN-1991
;       APPLICATION NUMBER:  US 07/666,056
;       FILING DATE:  07-MAR-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Frommer, William S.
;       REGISTRATION NUMBER:  25,506
;       REFERENCE/DOCKET NUMBER:  454310-2291
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 840-3333
;       TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  214:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  389 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  not relevant
US-08-036-218-214

  Query Match             96.9%;  Score 1870;  DB 4;  Length 389;
  Best Local Similarity   98.0%;  Pred. No. 1.8e-136;
  Matches  335;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||||||:|||||||||||||:||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||| ||||||| ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347
```

```
Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 6
US-08-036-218-217
; Sequence 217, Application US/08036218
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Perkus, Marion E.
; APPLICANT: Taylor, Jill
; APPLICANT: Tartaglia, James
; APPLICANT: Norton, Elizabeth K.
; APPLICANT: Riviere, Michel
; APPLICANT: de Taisne, Charles
; APPLICANT: Limbach, Keith J.
; APPLICANT: Johnson, Gerard P.
; APPLICANT: Pincus, Steven E.
; APPLICANT: Cox, William I.
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE STRAIN
; NUMBER OF SEQUENCES: 221
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William S. Frommer c/o
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United State of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,218
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/713,967
; FILING DATE: 11-JUN-1991
; APPLICATION NUMBER: US 07/666,056
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2291
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 217:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
US-08-036-218-217

Query Match 96.9%; Score 1870; DB 4; Length 389;
Best Local Similarity 98.0%; Pred. No. 1.8e-136;
Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||||:||||||||||||:||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||| ||||||| ||:||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 7
US-08-360-107-116
; Sequence 116, Application US/08360107
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 137
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,107
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-360-107-116

Query Match 96.8%; Score 1868; DB 7; Length 389;
Best Local Similarity 98.2%; Pred. No. 2.6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
```

```
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185

Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245

Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305

Db    288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347

Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 8
US-08-470-896-106
; Sequence 106, Application US/08470896
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,896
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-470-896-106

Query Match 96.8%; Score 1868; DB 8; Length 389;
Best Local Similarity 98.2%; Pred. No. 2.6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65

Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125

Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185

Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245

Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305

Db    288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347

Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 9
US-08-471-913-106
; Sequence 106, Application US/08471913
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,913
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-471-913-106

Query Match 96.8%; Score 1868; DB 8; Length 389;

```
  Best Local Similarity   98.2%;  Pred. No. 2.6e-136;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 10
US-08-475-668-106
; Sequence 106, Application US/08475668
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  COMPOSITIONS FOR INHIBITION OF
;    TITLE OF INVENTION:  MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA
;    TITLE OF INVENTION:  VIRUS TRANSMISSION
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,668
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-026
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-475-668-106

  Query Match             96.8%;  Score 1868;  DB 8;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 2.6e-136;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 11
US-08-484-223-106
; Sequence 106, Application US/08484223
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  COMPOSITIONS FOR INHIBITION OF
;    TITLE OF INVENTION:  MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
;    TITLE OF INVENTION:  TRANSMISSION
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,223
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-029
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
```

; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-223-106

Query Match 96.8%; Score 1868; DB 8; Length 389;
Best Local Similarity 98.2%; Pred. No. 2.6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||| || ||:||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 12
US-08-484-223A-106
; Sequence 106, Application US/08484223A
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-223A-106

Query Match 96.8%; Score 1868; DB 8; Length 389;
Best Local Similarity 98.2%; Pred. No. 2.6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||| || ||:||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 13
US-08-485-546-106
; Sequence 106, Application US/08485546
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,546
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.

```
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-028
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-485-546-106

  Query Match             96.8%;  Score 1868;  DB 8;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 2.6e-136;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 14
US-08-485-546A-106
; Sequence 106, Application US/08485546A
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  METHODS FOR INHIBITION OF MEMBRANE
;    TITLE OF INVENTION:  FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
;    TITLE OF INVENTION:  TRANSMISSION
;    NUMBER OF SEQUENCES:  214
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,546A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-028
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-485-546A-106

  Query Match             96.8%;  Score 1868;  DB 8;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 2.6e-136;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 15
US-08-485-551-106
; Sequence 106, Application US/08485551
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  METHODS FOR INHIBITION OF MEMBRANE
;    TITLE OF INVENTION:  FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
;    TITLE OF INVENTION:  TRANSMISSION
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/485,551
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Coruzzi, Laura A.
;       REGISTRATION NUMBER:  30,742
;       REFERENCE/DOCKET NUMBER:  7872-023
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 790-9090
;       TELEFAX:  (212) 869-9741/8864
;       TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  106:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  389 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
US-08-485-551-106

  Query Match             96.8%;  Score 1868;  DB 8;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 2.6e-136;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
            |||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db       48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy       66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
            |||||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db      108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy      126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy      186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy      246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
            | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy      306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
            |||||||||||||||||||||||||||||||||||||||||||
Db      348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


Search completed:  February  3, 2003, 09:34:07
Job time :  250.857 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:25:53 ; Search time 12.6951 Seconds
(without alignments)
551.547 Million cell updates/sec

Title: US-09-890-752A-1
Perfect score: 1929
Sequence: 1 MGRGDGAGAFGLGFTPPHGG.....SILSPFLPLLPIFFCLWVYI 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1866	96.7	389	9	US-09-821-877-2	Sequence 2, Appli
2	1721	89.2	400	9	US-10-209-264-3	Sequence 3, Appli
3	1534	79.5	281	9	US-09-247-890-12	Sequence 12, Appl
4	1437	74.5	281	9	US-09-247-890-10	Sequence 10, Appl
5	1259	65.3	226	9	US-10-104-966-10	Sequence 10, Appl
6	1259	65.3	226	10	US-09-929-955-10	Sequence 10, Appl
7	1259	65.3	229	9	US-09-821-877-8	Sequence 8, Appli
8	1252	64.9	226	10	US-09-812-862-14	Sequence 14, Appl
9	1231	63.8	397	10	US-09-812-862-6	Sequence 6, Appli
10	997	51.7	351	10	US-09-812-862-4	Sequence 4, Appli
11	834.5	43.3	282	9	US-09-247-890-16	Sequence 16, Appl
12	652	33.8	289	10	US-09-812-862-8	Sequence 8, Appli
13	649	33.6	346	10	US-09-812-862-2	Sequence 2, Appli
14	613	31.8	174	10	US-09-818-066-55	Sequence 55, Appl
15	604	31.3	174	10	US-09-818-066-54	Sequence 54, Appl
16	579	30.0	174	10	US-09-818-066-56	Sequence 56, Appl
17	571	29.6	174	10	US-09-818-066-57	Sequence 57, Appl
18	565	29.3	174	10	US-09-818-066-46	Sequence 46, Appl
19	561	29.1	174	10	US-09-818-066-48	Sequence 48, Appl
20	554	28.7	174	10	US-09-818-066-53	Sequence 53, Appl
21	552	28.6	174	10	US-09-818-066-45	Sequence 45, Appl
22	552	28.6	174	10	US-09-818-066-62	Sequence 62, Appl
23	551	28.6	174	10	US-09-818-066-42	Sequence 42, Appl
24	550	28.5	174	10	US-09-818-066-59	Sequence 59, Appl
25	545	28.3	174	10	US-09-818-066-43	Sequence 43, Appl
26	544	28.2	174	10	US-09-818-066-47	Sequence 47, Appl
27	541	28.0	170	10	US-09-818-066-49	Sequence 49, Appl
28	539	27.9	174	10	US-09-818-066-37	Sequence 37, Appl
29	539	27.9	174	10	US-09-818-066-44	Sequence 44, Appl
30	538	27.9	174	10	US-09-818-066-52	Sequence 52, Appl
31	533	27.6	174	10	US-09-818-066-58	Sequence 58, Appl
32	529	27.4	174	10	US-09-818-066-38	Sequence 38, Appl
33	528	27.4	174	10	US-09-818-066-50	Sequence 50, Appl
34	528	27.4	174	10	US-09-818-066-51	Sequence 51, Appl
35	525	27.2	174	10	US-09-818-066-40	Sequence 40, Appl
36	525	27.2	174	10	US-09-818-066-41	Sequence 41, Appl
37	522	27.1	174	10	US-09-818-066-39	Sequence 39, Appl
38	496	25.7	174	10	US-09-818-066-35	Sequence 35, Appl
39	488	25.3	174	10	US-09-818-066-36	Sequence 36, Appl
40	480	24.9	174	10	US-09-818-066-60	Sequence 60, Appl
41	472	24.5	174	10	US-09-818-066-61	Sequence 61, Appl
42	353	18.3	61	9	US-09-821-877-5	Sequence 5, Appli
43	228	11.8	55	10	US-09-879-257A-45	Sequence 45, Appl
44	209	10.8	48	10	US-09-818-066-67	Sequence 67, Appl
45	208.5	10.8	420	10	US-09-812-862-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-821-877-2
; Sequence 2, Application US/09821877
; Patent No. US20020177124A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Coleman, Paul F.
; APPLICANT: Mushahwar, Isa K.
; TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant
; TITLE OF INVENTION: And Methods Of Detection Thereof
; FILE REFERENCE: 6794.US.01
; CURRENT APPLICATION NUMBER: US/09/821,877
; CURRENT FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Hepatitis B Virus
US-09-821-877-2

Query Match 96.7%; Score 1866; DB 9; Length 389;
Best Local Similarity 98.2%; Pred. No. 1.9e-132;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||||||||||||||||||||||||||| :||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGITQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPARNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRAC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
```

```
Db    288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          ||||||||||||||| |||||||:||||||||||||||||||
Db    348 GLSPTVWLSVIWMMLYWGPSLYTILSPFLPLLPIFFCLWVYI 389


RESULT 2
US-10-209-264-3
; Sequence 3, Application US/10209264
; Publication No. US20030003111A1
;   GENERAL INFORMATION:
;        APPLICANT: Oon, Chong Jin
;                   Lim, Gek Keow
;                   Zhao, Yi
;                   Chen, Wei Ning
;        TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
;                            USES THEREOF
;        NUMBER OF SEQUENCES: 11
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Ladas & Parry
;             STREET: 26 West 61 Street
;             CITY: New York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10023
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/209,264
;             FILING DATE: 31-Jul-2002
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/SG98/00046
;             FILING DATE: 19-JAN-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mass, Clifford J.
;             REGISTRATION NUMBER: 30,086
;             REFERENCE/DOCKET NUMBER: U-013109-7
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 708-1800
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 400 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-209-264-3

  Query Match             89.2%;  Score 1721;  DB 9;  Length 400;
  Best Local Similarity   90.1%;  Pred. No. 1.3e-121;
  Matches  308;  Conservative   10;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          | |||| |||||||||| |||||||| |:|| |||||||||||||||:|||||::|||
Db     59 GVGAFGPGFTPPHGGLLGGSPQAQGILTTVPAAPPPASTNRQSGRQPTPISPPLRDSHPQ 118

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          | |||||||| | |||||||||||||||||||||||| ||||||||| |||| ||||
Db    119 ATQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRTGDPAPNMENT 178

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||| |||||||||||||||||||||    | ||||||||||||||||
Db    179 TSGFLGPLLVLQAGFFSLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 238

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||| |||||| ||||||||||||||||||||||||||||||:||:|||||||:||
Db    239 CPPICPGYRWNCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 298


Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          |||||| | :|||||||||||||||||||||||| :||||||| ||||||||||||||||
Db    299 TTPAQGNSTFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFV 358

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||| |||:||||||||||||||||||||
Db    359 GLSPTVWLSVIWMMWYWGRSLYNILSPFLPLLPIFFCLWVYI 400


RESULT 3
US-09-247-890-12
; Sequence 12, Application US/09247890
; Publication No. US20020198162A1
; GENERAL INFORMATION:
;  APPLICANT: Punnonen, Juha
;  APPLICANT: Bass, Steven H.
;  APPLICANT: Whalen, Robert Gerald
;  APPLICANT: Howard, Russell
;  APPLICANT: Stemmer, Willem P.C.
;  APPLICANT: Maxygen, Inc.
;  TITLE OF INVENTION: Antigen Library Immunization
;  FILE REFERENCE: 018097-028710US
;  CURRENT APPLICATION NUMBER: US/09/247,890
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: US 60/074,294
;  EARLIER FILING DATE: 1998-02-11
;  EARLIER APPLICATION NUMBER: US 60/105,509
;  EARLIER FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-247-890-12

  Query Match             79.5%;  Score 1534;  DB 9;  Length 281;
  Best Local Similarity   98.2%;  Pred. No. 8.5e-108;
  Matches  276;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     67 MQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENIT 126
          |||||||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db      1 MQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENIT 60

Qy    127 SGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTSC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 SGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTSC 120

Qy    187 PPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTCT 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 PPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTCM 180

Qy    247 TPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFVG 306
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFVG 240

Qy    307 LSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    241 LSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 281


RESULT 4
US-09-247-890-10
; Sequence 10, Application US/09247890
; Publication No. US20020198162A1
; GENERAL INFORMATION:
;  APPLICANT: Punnonen, Juha
;  APPLICANT: Bass, Steven H.
;  APPLICANT: Whalen, Robert Gerald
;  APPLICANT: Howard, Russell
;  APPLICANT: Stemmer, Willem P.C.
;  APPLICANT: Maxygen, Inc.
```

```
;  TITLE OF INVENTION: Antigen Library Immunization
;  FILE REFERENCE: 018097-028710US
;  CURRENT APPLICATION NUMBER: US/09/247,890
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: US 60/074,294
;  EARLIER FILING DATE: 1998-02-11
;  EARLIER APPLICATION NUMBER: US 60/105,509
;  EARLIER FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-247-890-10

  Query Match             74.5%;  Score 1437;  DB 9;  Length 281;
  Best Local Similarity   91.5%;  Pred. No. 1.5e-100;
  Matches  257;  Conservative    6;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      67 MQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENIT 126
           |||||||||| | ||||||||||||||||||||||||| |||||||||| |||| |||| |
Db       1 MQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRTGDPAPNMENTT 60

Qy     127 SGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTSC 186
           |||||||||||||||||||||||||||| ||||||||||     |||||||||||||||||
Db      61 SGFLGPLLVLQAGFFLLTRILTIPQSLHSWWTSLNFLGAAPTCLGQNSQSPTSNHSPTSC 120

Qy     187 PPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTCT 246
           || |||||||||||||||||||||||||||||||||||| |||:| :|||||||:|||
Db     121 PPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLSVCPLLPRTSTTSTGPCKTCT 180

Qy     247 TPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFVG 306
           |||||||:|| |||||||||||||||||||| :|||||||  ||||||||||||||||||
Db     181 IPAQGTSMFPSSCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFVG 240

Qy     307 LSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||:||||||||||||||||||||||
Db     241 LSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 281


RESULT 5
US-10-104-966-10
; Sequence 10, Application US/10104966
; Patent No. US20020155124A1
; GENERAL INFORMATION:
;  APPLICANT: Matti Sallberg
;  APPLICANT:  Catharina Hultgren
;  TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: TRIPEP.23AUSC1
;  CURRENT APPLICATION NUMBER: US/10/104,966
;  CURRENT FILING DATE:  2002-03-22
;  PRIOR APPLICATION NUMBER: 09/705,547
;  PRIOR FILING DATE: 2000-11-03
;  PRIOR APPLICATION NUMBER: 60/229,175
;  PRIOR FILING DATE: 2000-08-29
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 226
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hepatitis B virus S antigen (HBsAg) sequence
US-10-104-966-10

  Query Match             65.3%;  Score 1259;  DB 9;  Length 226;
  Best Local Similarity   99.1%;  Pred. No. 2.4e-87;
  Matches  224;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     122 MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNH 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNH 60

Qy     182 SPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGP 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGP 120

Qy     242 CRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFV 301
           |||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFV 180

Qy     302 QWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     181 QWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 226


RESULT 6
US-09-929-955-10
; Sequence 10, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
;  APPLICANT: Matti Sallberg
;  APPLICANT:  Catharina Hultgren
;  TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: TRIPEP.23AUS2
;  CURRENT APPLICATION NUMBER: US/09/929,955
;  CURRENT FILING DATE:  2001-08-15
;  PRIOR APPLICATION NUMBER: 09/705,547
;  PRIOR FILING DATE: 2000-11-03
;  PRIOR APPLICATION NUMBER: 60/229,175
;  PRIOR FILING DATE: 2000-08-29
;  PRIOR APPLICATION NUMBER: 60/225,767
;  PRIOR FILING DATE: 2000-08-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 226
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hepatitis B virus S antigen (HBsAg) sequence
US-09-929-955-10

  Query Match             65.3%;  Score 1259;  DB 10;  Length 226;
  Best Local Similarity   99.1%;  Pred. No. 2.4e-87;
  Matches  224;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     122 MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNH 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNH 60

Qy     182 SPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGP 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGP 120

Qy     242 CRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFV 301
           |||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFV 180

Qy     302 QWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     181 QWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 226


RESULT 7
US-09-821-877-8
; Sequence 8, Application US/09821877
; Patent No. US20020177124A1
; GENERAL INFORMATION:
;  APPLICANT: Abbott Laboratories
;  APPLICANT:  Coleman, Paul F.
;  APPLICANT:  Mushahwar, Isa K.
;  TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant
```

```
;  TITLE OF INVENTION:  And Methods Of Detection Thereof
;  FILE REFERENCE: 6794.US.01
;  CURRENT APPLICATION NUMBER: US/09/821,877
;  CURRENT FILING DATE:  2000-03-03
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;    LENGTH: 229
;    TYPE: PRT
;    ORGANISM: Hepatitis B Virus
;    FEATURE:
;    NAME/KEY: VARIANT
;    LOCATION: (126)...(126)
;    OTHER INFORMATION: Xaa = A or T at position 126
;    NAME/KEY: VARIANT
;    LOCATION: (202)...(202)
;    OTHER INFORMATION: Xaa = L or W at position 202
;    NAME/KEY: VARIANT
;    LOCATION: (210)...(210)
;    OTHER INFORMATION: Xaa = T or S at position 210
US-09-821-877-8

  Query Match             65.3%;  Score 1259;  DB 9;  Length 229;
  Best Local Similarity   98.3%;  Pred. No. 2.5e-87;
  Matches  225;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     119 ALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPT 178
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ARNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPT 60

Qy     179 SNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTS 238
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTS 120

Qy     239 TGPCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLV 298
           ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGPCRXCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLV 180

Qy     299 PFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           ||||||||||||||||||| ||||||| ||||||||||||||||||||
Db     181 PFVQWFVGLSPTVWLSVIWMMXYWGPSLYXILSPFLPLLPIFFCLWVYI 229


RESULT 8
US-09-812-862-14
; Sequence 14, Application US/09812862
; Patent No. US20020035081A1
;   GENERAL INFORMATION:
;        APPLICANT: Wands, Jack R.
;                   Scaglioni, Pier Paolo
;                   Melegari, Margherita
;        TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
;        NUMBER OF SEQUENCES: 14
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/812,862
;             FILING DATE: 20-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/667,073
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/492,489
;             FILING DATE: 20-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Clark, Paul T.
;             REGISTRATION NUMBER: 30,162
;             REFERENCE/DOCKET NUMBER: 00786/282001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617/542-5070
;             TELEFAX: 617/542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 14:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 226 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-812-862-14

  Query Match             64.9%;  Score 1252;  DB 10;  Length 226;
  Best Local Similarity   98.7%;  Pred. No. 8.1e-87;
  Matches  223;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     122 MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNH 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNH 60

Qy     182 SPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGP 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGP 120

Qy     242 CRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFV 301
           |||| | ||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     121 CRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASAPFSWLSLLVPFV 180

Qy     302 QWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           ||||||||||||||||||||||||||||||||||||||||||||||
Db     181 QWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 226


RESULT 9
US-09-812-862-6
; Sequence 6, Application US/09812862
; Patent No. US20020035081A1
;   GENERAL INFORMATION:
;        APPLICANT: Wands, Jack R.
;                   Scaglioni, Pier Paolo
;                   Melegari, Margherita
;        TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
;        NUMBER OF SEQUENCES: 14
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/812,862
;             FILING DATE: 20-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/667,073
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/492,489
;             FILING DATE: 20-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Clark, Paul T.
;             REGISTRATION NUMBER: 30,162
```

```
;           REFERENCE/DOCKET NUMBER: 00786/282001
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 617/542-5070
;           TELEFAX: 617/542-8906
;           TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 6:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 397 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-812-862-6

  Query Match             63.8%;  Score 1231;  DB 10;  Length 397;
  Best Local Similarity   69.7%;  Pred. No. 5.6e-85;
  Matches  237;  Conservative    8;  Mismatches   31;  Indels   64;  Gaps    5;

Qy       9 AFGLGF-TPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQAM 67
          :||:   |||      : |    || ||    |  :  |: || |   :|  |     |:
Db     121 SFGVWIRTPP-----AYRPPNAPILSTL----PETTVVRRRGRSPRRRTPSPRRRRSQS- 170

Qy      68 QWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENITS 127
                    || |                                              :
Db     171 ------------PRRRR-----------------------------------------SQ 177

Qy     128 GFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTSCP 187
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     178 SRLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTSCP 237

Qy     188 PTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTCTT 247
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     238 PTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTCMT 297

Qy     248 PAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFVGL 307
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     298 TAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFVGL 357

Qy     308 SPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||
Db     358 SPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 397


RESULT 10
US-09-812-862-4
; Sequence 4, Application US/09812862
; Patent No. US20020035081A1
;   GENERAL INFORMATION:
;       APPLICANT: Wands, Jack R.
;                  Scaglioni, Pier Paolo
;                  Melegari, Margherita
;       TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
;       NUMBER OF SEQUENCES: 14
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: Fish & Richardson P.C.
;           STREET: 225 Franklin Street
;           CITY: Boston
;           STATE: MA
;           COUNTRY: USA
;           ZIP: 02110-2804
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/812,862
;           FILING DATE: 20-Mar-2001
;           CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 08/667,073
;           FILING DATE: <Unknown>
;           APPLICATION NUMBER: 08/492,489
;           FILING DATE: 20-JUN-1995
;       ATTORNEY/AGENT INFORMATION:
;           NAME: Clark, Paul T.
;           REGISTRATION NUMBER: 30,162
;           REFERENCE/DOCKET NUMBER: 00786/282001
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 617/542-5070
;           TELEFAX: 617/542-8906
;           TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 351 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-812-862-4

  Query Match             51.7%;  Score 997;  DB 10;  Length 351;
  Best Local Similarity   98.9%;  Pred. No. 1.5e-67;
  Matches  174;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     172 QNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLI 231
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     176 QNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLI 235

Qy     232 PGSSTTSTGPCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARF 291
           |||||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db     236 PGSSTTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARF 295

Qy     292 SWLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     296 SWLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 351


RESULT 11
US-09-247-890-16
; Sequence 16, Application US/09247890
; Publication No. US20020198162A1
; GENERAL INFORMATION:
;  APPLICANT: Punnonen, Juha
;  APPLICANT: Bass, Steven H.
;  APPLICANT: Whalen, Robert Gerald
;  APPLICANT: Howard, Russell
;  APPLICANT: Stemmer, Willem P.C.
;  APPLICANT: Maxygen, Inc.
;  TITLE OF INVENTION: Antigen Library Immunization
;  FILE REFERENCE: 018097-028710US
;  CURRENT APPLICATION NUMBER: US/09/247,890
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: US 60/074,294
;  EARLIER FILING DATE: 1998-02-11
;  EARLIER APPLICATION NUMBER: US 60/105,509
;  EARLIER FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
;   LENGTH: 282
;   TYPE: PRT
;   ORGANISM: Woodchuck hepatitis B virus
US-09-247-890-16

  Query Match             43.3%;  Score 834.5;  DB 9;  Length 282;
  Best Local Similarity   56.8%;  Pred. No. 1.7e-55;
  Matches  159;  Conservative   29;  Mismatches   85;  Indels    7;  Gaps    2;

Qy      73 TFH-----QTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENITS 127
           |||     |:|          | |     |::|   ||| | |  |  |||||: |   |
Db       5 TFHLQGFVDGLRDLTTTERQHNAYGDPFTTLSPAVPTVSTILSPPSTTGDPALSPEMSPS 64

Qy     128 GFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTSCP 187
           || |  ||  :|| |:|||| |:|| ||||||:| ||   | |||||  |  | ||||||
Db      65 SLLGLLAGLQVVYFLWTKILTIAQNLDWWWTSLSFPGGIPECTGQNSQFQTCKHLPTSCP 124
```

```
Qy     188 PTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTCTT 247
           ||| |:||| |||||||:| :|||||||||||||||::|::||||| | : ||    || ||
Db     125 PTCNGFRWMYLRRFIIYLLVLLLCLIFLLVLLDWKGLIPVCPLQPTTETTVN--CRQCTI 182

Qy     248 PAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFVGL 307
             ||      | ||| ||: ||||| ||||||| | :||||| ||||||:|||| :||  |:
Db     183 SAQDMYTPPYCCCLKPTAGNCTCWPIPSSWALGNYLWEWALARFSWLNLLVPLLQWLGGI 242

Qy     308 SPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIPFCLWVYI 347
           |   |  :|||:|:|||:| ||| ||:|:  :|| :||||
Db     243 SLIAWFLLIWMIWFWGPALLSILPPFIPIFVLFFLIWVYI 282


RESULT 12
US-09-812-862-8
; Sequence 8, Application US/09812862
; Patent No. US20020035081A1
;   GENERAL INFORMATION:
;        APPLICANT: Wands, Jack R.
;                   Scaglioni, Pier Paolo
;                   Melegari, Margherita
;        TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
;        NUMBER OF SEQUENCES: 14
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/812,862
;             FILING DATE: 20-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/667,073
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/492,489
;             FILING DATE: 20-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Clark, Paul T.
;             REGISTRATION NUMBER: 30,162
;             REFERENCE/DOCKET NUMBER: 00786/282001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617/542-5070
;             TELEFAX: 617/542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 289 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-812-862-8

  Query Match             33.8%;  Score 652;  DB 10;  Length 289;
  Best Local Similarity   98.2%;  Pred. No. 7.5e-42;
  Matches  112;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     234 SSTTSTGPCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSW 293
           ||||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db     176 SSTTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSW 235

Qy     294 LSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIPFCLWVYI 347
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     236 LSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIPFCLWVYI 289


RESULT 13
US-09-812-862-2
; Sequence 2, Application US/09812862
; Patent No. US20020035081A1
;   GENERAL INFORMATION:
;        APPLICANT: Wands, Jack R.
;                   Scaglioni, Pier Paolo
;                   Melegari, Margherita
;        TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
;        NUMBER OF SEQUENCES: 14
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/812,862
;             FILING DATE: 20-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/667,073
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/492,489
;             FILING DATE: 20-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Clark, Paul T.
;             REGISTRATION NUMBER: 30,162
;             REFERENCE/DOCKET NUMBER: 00786/282001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617/542-5070
;             TELEFAX: 617/542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 346 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-812-862-2

  Query Match             33.6%;  Score 649;  DB 10;  Length 346;
  Best Local Similarity   63.1%;  Pred. No. 1.5e-41;
  Matches  111;  Conservative   21;  Mismatches   42;  Indels    2;  Gaps    1;

Qy     172 QNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLI 231
           ||||  |  | |||||||| |:||| |||||:| :|||||||||||||::|::|||||
Db     173 QNSQFQTCKHLPTSCPPTCNGFRWMYLRRFIIYLLVLLLCLIFLLVLLDWKGLIPVCPLQ 232

Qy     232 PGSSTTSTGPCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARF 291
           | : ||    || ||  ||      | ||| ||: ||||| ||||||| | :||||| ||
Db     233 PTTETTVN--CRQCTISAQDMYTPPYCCCLKPTAGNCTCWPIPSSWALGNYLWEWALARL 290

Qy     292 SWLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIPFCLWVYI 347
           |||:|||| :||  |:|   |  :|||:|:|||:| ||| ||:|:  :|| :||||
Db     291 SWLNLLVPLLQWLGGISLIAWFLLIWMIWFWGPALLSILPPFIPIFVLFFLIWVYI 346


RESULT 14
US-09-818-066-55
; Sequence 55, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;        APPLICANT: Shuping Tong et al.
```

```
;       TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;       NUMBER OF SEQUENCES: 75
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fish & Richardson P.C.
;            STREET: 225 Franklin Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02110-2804
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/818,066
;            FILING DATE: 27-Mar-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/683,262
;            FILING DATE: 18-JUL-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Fraser, Janis K.
;            REGISTRATION NUMBER: 31,819
;            REFERENCE/DOCKET NUMBER: 00786/287002
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 542-5070
;            TELEFAX: (617) 542-8906
;            TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 55:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 174 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-818-066-55

  Query Match           31.8%;  Score 613;  DB 10;  Length 174;
  Best Local Similarity  96.6%;  Pred. No. 3.5e-39;
  Matches  112;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 GAGAPGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||||||||||||||||||||||||||::||||||||||||||||||||||||||| ||||
Db      59 GAGAFGLGFTPPHGGLLGWSPQAQGIIQTLPANPPPASTNRQSGRQPTPLSPPLRTTHPQ 118

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALN 121
           ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     119 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPALN 174


RESULT 15
US-09-818-066-54
; Sequence 54, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;       APPLICANT: Shuping Tong et al.
;       TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;       NUMBER OF SEQUENCES: 75
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fish & Richardson P.C.
;            STREET: 225 Franklin Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02110-2804
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/818,066
;            FILING DATE: 27-Mar-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/683,262
;            FILING DATE: 18-JUL-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Fraser, Janis K.
;            REGISTRATION NUMBER: 31,819
;            REFERENCE/DOCKET NUMBER: 00786/287002
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 542-5070
;            TELEFAX: (617) 542-8906
;            TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 54:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 174 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-818-066-54

  Query Match           31.3%;  Score 604;  DB 10;  Length 174;
  Best Local Similarity  96.6%;  Pred. No. 1.7e-38;
  Matches  112;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6 GAGAPGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      59 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 118

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALN 121
           |||||||||||||||||||||| |||||||||||||||||||||||||  |||||
Db     119 AMQWNSTTFHQTLQDPRVRGLYLPAGGSSSGTVNPVPTTVSPISSIFSRTIDPALN 174


Search completed: February  3, 2003, 09:35:27
Job time : 13.6951 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:22:17 ; Search time 21.1585 Seconds
(without alignments)
482.536 Million cell updates/sec

Title: US-09-890-752A-1
Perfect score: 1929
Sequence: 1 MGRGDGAGAFGLGFTPPHGG.....SILSPFLPLLLPIFFCLWVYI 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1870	96.9	389	1	US-08-105-483-216	Sequence 216, App
2	1870	96.9	389	1	US-08-105-483-219	Sequence 219, App
3	1870	96.9	389	1	US-08-709-209-216	Sequence 216, App
4	1870	96.9	389	1	US-08-709-209-219	Sequence 219, App
5	1870	96.9	389	1	US-08-458-101-216	Sequence 216, App
6	1870	96.9	389	1	US-08-458-101-219	Sequence 219, App
7	1868	96.8	389	3	US-08-486-099-106	Sequence 106, App
8	1868	96.8	389	3	US-08-360-107A-116	Sequence 116, App
9	1868	96.8	389	3	US-08-484-223B-106	Sequence 106, App
10	1868	96.8	389	3	US-08-919-597-106	Sequence 106, App
11	1868	96.8	389	3	US-08-475-668A-106	Sequence 106, App
12	1868	96.8	389	3	US-08-485-551A-106	Sequence 106, App
13	1868	96.8	389	3	US-08-471-913A-106	Sequence 106, App
14	1868	96.8	389	4	US-08-485-264A-106	Sequence 106, App
15	1868	96.8	389	4	US-08-474-349A-106	Sequence 106, App
16	1648.5	85.5	395	6	5196194-18	Patent No. 5196194
17	1539	79.8	281	1	US-08-105-483-214	Sequence 214, App
18	1539	79.8	281	1	US-08-709-209-214	Sequence 214, App
19	1539	79.8	281	1	US-08-458-101-214	Sequence 214, App
20	1431	74.2	277	6	5164485-2	Patent No. 5164485
21	1252	64.9	226	5	PCT-US96-10602-14	Sequence 14, Appl
22	1240	64.3	236	1	US-08-378-011A-1	Sequence 1, Appli
23	1233.5	63.9	423	2	US-08-760-797A-1	Sequence 1, Appli
24	1233.5	63.9	424	4	US-08-932-929B-1	Sequence 1, Appli
25	1233	63.9	424	2	US-08-760-797A-3	Sequence 3, Appli
26	1233	63.9	424	4	US-08-932-929B-3	Sequence 3, Appli
27	1231	63.8	397	5	PCT-US96-10602-6	Sequence 6, Appli
28	1217	63.1	226	6	5196194-21	Patent No. 5196194
29	1217	63.1	226	6	5436139-5	Patent No. 5436139
30	1212	62.8	225	6	5436139-4	Patent No. 5436139
31	1206	62.5	226	1	US-08-378-011A-3	Sequence 3, Appli
32	1197	62.1	226	6	5198348-1	Patent No. 5198348
33	1160	60.1	228	1	US-08-447-591-2	Sequence 2, Appli
34	1160	60.1	228	1	US-08-450-943-2	Sequence 2, Appli
35	1160	60.1	228	1	US-08-059-031-2	Sequence 2, Appli
36	1160	60.1	228	2	US-08-450-942-2	Sequence 2, Appli
37	1160	60.1	228	5	PCT-US94-05090-2	Sequence 2, Appli
38	1142	59.2	228	1	US-08-447-591-3	Sequence 3, Appli
39	1142	59.2	228	1	US-08-450-943-3	Sequence 3, Appli
40	1142	59.2	228	1	US-08-059-031-3	Sequence 3, Appli
41	1142	59.2	228	2	US-08-450-942-3	Sequence 3, Appli
42	1142	59.2	228	5	PCT-US94-05090-3	Sequence 3, Appli
43	1114	57.8	228	1	US-08-447-591-4	Sequence 4, Appli
44	1114	57.8	228	1	US-08-450-943-4	Sequence 4, Appli
45	1114	57.8	228	1	US-08-059-031-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-105-483-216
; Sequence 216, Application US/08105483
; Patent No. 5494807
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/105,483
; FILING DATE: 12-AUG-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 216:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-105-483-216

Query Match 96.9%; Score 1870; DB 1; Length 389;
Best Local Similarity 98.0%; Pred. No. 3.4e-157;
Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65

```
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           ||||||||||||||||||||||||||||||| ||||||| ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 2
US-08-105-483-219
; Sequence 219, Application US/08105483
; Patent No. 5494807
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/105,483
;      FILING DATE:  12-AUG-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-105-483-219

  Query Match             96.9%;  Score 1870;  DB 1;  Length 389;
  Best Local Similarity   98.0%;  Pred. No. 3.4e-157;
  Matches  335;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           |||||||||||||||||||||||||||||:||||||||||:||||||||||||||||||
```

```
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           ||||||||||||||||||||||||||||||| ||||||| ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 3
US-08-709-209-216
; Sequence 216, Application US/08709209
; Patent No. 5762938
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/709,209
;      FILING DATE:  21-AUG-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/105,483
;      FILING DATE:  12-AUG-1993
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  216:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-709-209-216

  Query Match             96.9%;  Score 1870;  DB 1;  Length 389;
  Best Local Similarity   98.0%;  Pred. No. 3.4e-157;
  Matches  335;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 4
US-08-709-209-219
; Sequence 219, Application US/08709209
; Patent No. 5762938
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,209
; FILING DATE: 21-AUG-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/105,483
; FILING DATE: 12-AUG-1993
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-709-209-219

Query Match 96.9%; Score 1870; DB 1; Length 389;
Best Local Similarity 98.0%; Pred. No. 3.4e-157;

Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 5
US-08-458-101-216
; Sequence 216, Application US/08458101
; Patent No. 5766599
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Perkus, Marion E.
; APPLICANT: Taylor, Jill
; APPLICANT: Tartaglia, James
; APPLICANT: No. 5766599ton, Elizabeth K.
; APPLICANT: Riviere, Michel
; APPLICANT: de Taisne, Charles
; APPLICANT: Limbach, Keith J.
; APPLICANT: Johnson, Gerard P.
; APPLICANT: Pincus, Steven E.
; APPLICANT: Cox, William I.
; APPLICANT: Audonnet, Jean-Christophe Francis
; APPLICANT: Gettig, Russell Robert
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,101
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2740
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 216:

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-458-101-216

  Query Match             96.9%;  Score 1870;  DB 1;  Length 389;
  Best Local Similarity   98.0%;  Pred. No. 3.4e-157;
  Matches  335;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:|||||||||||||:||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||| ||||||| ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 6
US-08-458-101-219
; Sequence 219, Application US/08458101
; Patent No. 5766599
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Perkus, Marion E.
;    APPLICANT:  Taylor, Jill
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  No. 5766599ton, Elizabeth K.
;    APPLICANT:  Riviere, Michel
;    APPLICANT:  de Taisne, Charles
;    APPLICANT:  Limbach, Keith J.
;    APPLICANT:  Johnson, Gerard P.
;    APPLICANT:  Pincus, Steven E.
;    APPLICANT:  Cox, William I.
;    APPLICANT:  Audonnet, Jean-Christophe Francis
;    APPLICANT:  Gettig, Russell Robert
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,101
;      FILING DATE:  01-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2740
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-458-101-219

  Query Match             96.9%;  Score 1870;  DB 1;  Length 389;
  Best Local Similarity   98.0%;  Pred. No. 3.4e-157;
  Matches  335;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:|||||||||||||:||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||| ||||||| ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 7
US-08-486-099-106
; Sequence 106, Application US/08486099
; Patent No. 6013263
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  COMPOSITIONS FOR INHIBITION OF
;    TITLE OF INVENTION:  MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
;    TITLE OF INVENTION:  B VIRUS TRANSMISSION
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,099
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-031
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-486-099-106

  Query Match             96.8%;  Score 1868;  DB 3;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 5.2e-157;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 8
US-08-360-107A-116
; Sequence 116, Application US/08360107A
; Patent No. 6017536
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR INHIBITION
;    TITLE OF INVENTION:  OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
;    TITLE OF INVENTION:  TRANSMISSION
;    NUMBER OF SEQUENCES:  149
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711

;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/360,107A
;      FILING DATE:  20-DEC-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-013
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  116:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-360-107A-116

  Query Match             96.8%;  Score 1868;  DB 3;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 5.2e-157;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 9
US-08-484-223B-106
; Sequence 106, Application US/08484223B
; Patent No. 6020459
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  COMPOSITIONS FOR INHIBITION OF
;    TITLE OF INVENTION:  MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
;    TITLE OF INVENTION:  TRANSMISSION
;    NUMBER OF SEQUENCES:  245
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
```

```
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,223B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-029
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-484-223B-106

  Query Match             96.8%;  Score 1868;  DB 3;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 5.2e-157;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 10
US-08-919-597-106
; Sequence 106, Application US/08919597
; Patent No. 6054265
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR INHIBITION
;    TITLE OF INVENTION:  OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
;    TITLE OF INVENTION:  TRANSMISSION
;    NUMBER OF SEQUENCES:  273
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/919,597
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/470,896
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-919-597-106

  Query Match             96.8%;  Score 1868;  DB 3;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 5.2e-157;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 11
US-08-475-668A-106
; Sequence 106, Application US/08475668A
; Patent No. 6060065
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    TITLE OF INVENTION:  COMPOSITIONS FOR INHIBITION OF MEMBRANE
;    TITLE OF INVENTION:  FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
;    TITLE OF INVENTION:  TRANSMISSION
;    NUMBER OF SEQUENCES:  211
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,668A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-026
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-475-668A-106

  Query Match             96.8%;  Score 1868;  DB 3;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 5.2e-157;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAPGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAPGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 12
US-08-485-551A-106
; Sequence 106, Application US/08485551A
; Patent No. 6068973
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  METHODS FOR INHIBITION OF MEMBRANE
;    TITLE OF INVENTION:  FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
;    TITLE OF INVENTION:  TRANSMISSION
;    NUMBER OF SEQUENCES:  211
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,551A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-485-551A-106

  Query Match             96.8%;  Score 1868;  DB 3;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 5.2e-157;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAPGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAPGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

```
RESULT 13
US-08-471-913A-106
; Sequence 106, Application US/08471913A
; Patent No. 6093794
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  COMPOSITIONS FOR INHIBITION OF MEMBRANE
;    TITLE OF INVENTION:  FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
;    TITLE OF INVENTION:  TRANSMISSION
;    NUMBER OF SEQUENCES:  214
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,913A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-030
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-471-913A-106

  Query Match             96.8%;  Score 1868;  DB 3;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 5.2e-157;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||| || ||:||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 14
US-08-485-264A-106
; Sequence 106, Application US/08485264A
; Patent No. 6228983
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  COMPOSITIONS FOR INHIBITION OF
;    TITLE OF INVENTION:  MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
;    TITLE OF INVENTION:  RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
;    NUMBER OF SEQUENCES:  232
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,264A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-021
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-485-264A-106

  Query Match             96.8%;  Score 1868;  DB 4;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 5.2e-157;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||| || ||:||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


RESULT 15
US-08-474-349A-106
; Sequence 106, Application US/08474349A
; Patent No. 6333395
;  GENERAL INFORMATION:
;    APPLICANT:  Bolognesi, Dani P.
;    APPLICANT:  Matthews, Thomas J.
;    APPLICANT:  Wild, Carl T.
;    APPLICANT:  Barney, Shawn O.
;    APPLICANT:  Lambert, Dennis M.
;    APPLICANT:  Petteway, Stephen R.
;    APPLICANT:  Langlois, Alphonse J.
;    TITLE OF INVENTION:  COMPOSITIONS FOR INHIBITION OF MEMBRANE
;    TITLE OF INVENTION:  FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
;    TITLE OF INVENTION:  VIRUS TRANSMISSION
;    NUMBER OF SEQUENCES:  517
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,349A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7872-024
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  389 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-474-349A-106

  Query Match             96.8%;  Score 1868;  DB 4;  Length 389;
  Best Local Similarity   98.2%;  Pred. No. 5.2e-157;
  Matches  336;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      6 GAGAPGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db     48 GAGAPGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||| || ||:||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389


Search completed: February  3, 2003, 09:27:09
Job time : 23.1585 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:15:41 ; Search time 55.6167 Seconds
(without alignments)
831.368 Million cell updates/sec

Title: US-09-890-752A-1
Perfect score: 1929
Sequence: 1 MGRGDGAGAFGLGFTPPHGG.....SILSPFLPLLPIFFCLWVYI 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1929	100.0	347	21	AAB10596	HBV fusion protein
2	1879	97.4	389	22	AAG66931	HBV genotype D pre
3	1870	96.9	389	13	AAR27472	lpsAg protein. Sy
4	1870	96.9	389	13	AAR27474	I3L promoter/S12/c
5	1868	96.8	389	22	AAG62931	Amino acid sequenc
6	1803.5	93.5	393	22	AAU14097	Hepatitis B virus
7	1783	92.4	399	22	AAG66932	HBV genotype E pre
8	1756	91.0	400	17	AAR93801	Hepatitis B virus
9	1752	90.8	389	7	AAP60794	Adr-type hepatitis
10	1745	90.5	400	22	AAG66930	HBV genotype C pre
11	1743	90.4	399	22	AAG66919	HBV HBPol protein.
12	1736	90.0	383	9	AAP80416	Sequence of adr ty
13	1727	89.5	389	15	AAR55287	Deduced sequence o
14	1721	89.2	400	21	AAY54045	Amino acid seqeunc
15	1713	88.8	388	12	AAR10851	Modified hepatitis
16	1713	88.8	389	23	AAM52682	Hepatitis B virus
17	1704	88.3	400	21	AAY44349	Human hepatitis B
18	1700	88.1	400	22	AAG66928	HBV genotype A pre
19	1686.5	87.4	387	12	AAR10852	Modified hepatitis
20	1684	87.3	400	22	AAG66929	HBV genotype B pre
21	1587	82.3	400	22	AAG66933	HBV genotype F pre
22	1539	79.8	281	13	AAR27471	spsAg protein. Sy
23	1534	79.5	281	20	AAY32835	HBsAg PreS2-S regi
24	1498.5	77.7	276	19	AAW44780	Modified hepatitis
25	1469	76.2	281	7	AAP60163	Subtype adr hepati
26	1469	76.2	281	7	AAP60560	Hepatitis B virus
27	1469	76.2	281	7	AAP60617	C-terminal defecti
28	1465	75.9	281	8	AAP70294	Subtype adr HBsAg
29	1460	75.7	281	15	AAR62870	Hepatitis B Virus
30	1460	75.7	281	22	AAM48419	Protein #3 used to
31	1444	74.9	281	22	AAE02620	HBsAg preS2-S doma
32	1437	74.5	281	20	AAY32834	HBsAg PreS2-S regi
33	1436	74.4	281	23	AAM52693	Hepatitis B virus
34	1428	74.0	281	15	AAR55281	Deduced sequence o
35	1418	73.5	281	7	AAP60561	Hepatitis B virus
36	1384	71.7	281	7	AAP60164	Subtype adw hepati
37	1279	66.3	249	21	AAY52545	Pan DR epitope/HBV
38	1265	65.6	226	22	AAG62930	Amino acid sequenc
39	1264	65.5	269	22	AAM48418	Protein #2 used to
40	1262.5	65.4	280	12	AAR10850	Modified hepatitis
41	1260	65.3	226	19	AAW62826	Hepatitis B virus
42	1259	65.3	226	23	AAE19897	Hepatitis B virus
43	1252	64.9	226	2	AAP10001	Sequence encoded b
44	1252	64.9	226	18	AAW09045	Hepatitis B virus
45	1249	64.7	226	23	ABB77764	Amino acid sequenc

ALIGNMENTS

RESULT 1
AAB10596
ID AAB10596 standard; Protein; 347 AA.
XX
AC AAB10596;
XX
DT 08-JAN-2001 (first entry)
XX
DE HBV fusion protein comprising LHB and RGD.
XX
KW Fusion protein; protein coat; virus-specific packaging signal; psi;
KW virus protein; cell permeability; cell-specific binding site; LHB;
KW large surface protein; core antigen; gene therapy.
XX
OS Hepatitis b virus.
OS Synthetic.
XX
PN WO200046376-A2.
XX
PD 10-AUG-2000.
XX
PF 04-FEB-2000; 2000WO-DE00363.
XX
PR 05-FEB-1999; 99DE-1004800.
XX
PA (HILD/) HILDT E.
XX
PI Hildt E, Hofschneider P;
XX
DR WPI; 2000-514959/46.
DR N-PSDB; AAA71734.
XX
PT Particle for cell-specific gene delivery, useful in gene therapy,

PT comprises nucleic acid in protein coat that includes a fusion protein
PT of viral protein, permeability peptide and cell-binding site -
XX
PS Claim 14; Fig 1; 34pp; German.
XX
CC This invention describes a novel particle (A), comprising a protein coat
CC with a fusion protein (FP), and, inside the coat, a nucleic acid (I)
CC including the sequence for a virus-specific packaging signal (psi) and a
CC structural gene. FP contains a virus protein (VP), a peptide (P) that
CC mediates cell permeability and a heterologous cell-specific binding site
CC (RGD). The invention also describes (1) producing (A) in which FP
CC contains an LHBs (large surface protein of hepatitis B virus (HBV)) and
CC RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),
CC (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector
CC containing the DNA of (d). The products of the invention are used in gene
CC therapy of cells and tissues, in vivo or ex vivo. This sequence
CC represents a fusion protein which is described in the method of the
CC invention.
XX
SQ Sequence 347 AA;

Query Match 100.0%; Score 1929; DB 21; Length 347;
Best Local Similarity 100.0%; Pred. No. 9.2e-143;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MGRGDGAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MGRGDGAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLR 60

Qy      61 NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPAL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPAL 120

Qy     121 NMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSN 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSN 180

Qy     181 HSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 HSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTG 240

Qy     241 PCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 PCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPF 300

Qy     301 VQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
```

RESULT 2
AAG66931
ID AAG66931 standard; Protein; 389 AA.
XX
AC AAG66931;
XX
DT 19-OCT-2001 (first entry)
XX
DE HBV genotype D preS1/preS2/HBsAg polypeptide.
XX
KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW HBsAg; antiviral; vaccine; genotype G; genotyping; HBcAg; HBeAg.
XX
OS Hepatitis B virus.
XX
PN WO200140279-A2.
XX
PD 07-JUN-2001.
XX
PF 20-NOV-2000; 2000WO-EP11526.
XX
PR 03-DEC-1999; 99EP-0870252.
PR 07-DEC-1999; 99US-0169287.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy -
XX
PS Example 3; Fig 6; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype
CC G. This genotype was found with a high prevalence in patients
CC chronically infected with HBV and residing in Europe and the USA. The
CC invention relates to a fully defined sequence of 3248 nucleotides as
CC given in specification, a sequence with 92% identity to the given
CC sequence, or sequence that is degenerate to the mentioned sequences.
CC These polynucleotides are useful for HBV genotyping. The proteins
CC encoded by the polynucleotides are useful for detecting antibodies in
CC a biological sample. Ligands that bind to the proteins and antibodies
CC directed against the proteins are useful for detecting the proteins
CC and for detecting HBcAg and HBeAg (precore precursor proteins). They
CC are also useful for preparing a vaccine or medicament for treating
CC HBV infections. The present sequence is provided in an amino acid
CC sequence alignment of the preS1, preS2 and HBsAg open reading frame
CC of the different HBV genotypes.
XX
SQ Sequence 389 AA;

Query Match 97.4%; Score 1879; DB 22; Length 389;
Best Local Similarity 98.5%; Pred. No. 8.4e-139;
Matches 337; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||::|||||||||||||||||||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGIIQTLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSAGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 3
AAR27472
ID AAR27472 standard; Protein; 389 AA.
XX
AC AAR27472;
XX
DT 24-FEB-1993 (first entry)
XX
DE lpsAg protein.
XX
KW Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;

KW vaccinia virus; hemorraghic region; u; promoter; NYVAC; recombinant;
KW HBV L; large pre-S antigen; lpsAg; fusion protein; pre-S region;
KW S12/core; S1; S2; Copenhagen vaccine strain; vaccinia virus;
KW virulence factor; deletion loci; recipient loci.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT region 1..108
FT /label S1
FT region 109..163
FT /label= S2
FT misc_RNA 164..389
FT /label= S
XX
PN WO9215672-A.
XX
PD 17-SEP-1992.
XX
PF 09-MAR-1992; 92WO-US01906.
XX
PR 07-MAR-1991; 91US-0666056.
PR 11-JUN-1991; 91US-0713967.
PR 06-MAR-1992; 92US-0847951.
XX
PA (VIRO-) VIROGENETICS CORP.
XX
PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP;
PI Limbach KJ, Norton EK, Paoletti E, Perkus ME, Pincus SE;
PI Riviere M, Tartaglia J, Taylor J;
XX
DR WPI; 1992-331718/40.
DR N-PSDB; AAQ29104.
XX
PT Vaccine comprises recombinant, attenuated pox-virus - use for
PT vaccinating against viral infections such as rabies, hepatitis B,
PT HIV, HSV, EBV, CMV, mumps etc.
XX
PS Disclosure; Fig 11; 456pp; English.
XX
CC The sequence given is encoded by an expression vector which comprises
CC the hepatitis B virus (HBV) L protein (large pre-S antigen, lpsAg)
CC gene linked to the cowpox hemorraghic region (u) promoter. This DNA
CC sequence was used in the construction of a NYVAC recombinant
CC expressing the HBV gene. Other HBV genes were also used in the
CC construction. These were HBV M protein (small pre-S antigen, spsAg)
CC and a fusion protein composed of the entire pre-S region (S12/core,
CC S1 + S2). Each of these gene sequences were inserted individually
CC into three different sites of NYVAC separated by from each other by
CC large regions of vaccinia DNA containing essential genes. NYVAC is a
CC Copenhagen vaccine strain of vaccinia virus which has been modified by
CC deletion of six non-essential regions of the genome encoding known or
CC potential virulence factors. The deletion loci were engineered as
CC recipient loci for the insertion of foreign genes. The spacing of the
CC three inserted sequences ensured that any recombination that did occur
CC would lead to disruption of the vaccinia genome and would cause
CC unviable vaccinia virus. See also AAQ35501-864.
XX
SQ Sequence 389 AA;

Query Match 96.9%; Score 1870; DB 13; Length 389;
Best Local Similarity 98.0%; Pred. No. 4.2e-138;
Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||:|||||||||||||:||||||||||||||||||
Db     48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||| ||||||| ||:||||||||||||||||||
Db    108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 4
AAR27474
ID AAR27474 standard; Protein; 389 AA.
XX
AC AAR27474;
XX
DT 24-FEB-1993 (first entry)
XX
DE I3L promoter/S12/core gene.
XX
KW Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
KW vaccinia virus; Amsacta moorei entomopoxvirus; AmEPV; 42 kD; promoter;
KW NYVAC; recombinant; HBV L; large pre-S antigen; lpsAg; fusion protein;
KW pre-S region; S12/core; S1; S2; Copenhagen vaccine strain;
KW vaccinia virus; virulence factor; deletion loci; recipient loci.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..108
FT /label= S1
FT Region 109..163
FT /label= S2
FT Region 164..389
FT /label= S
XX
PN WO9215672-A.
XX
PD 17-SEP-1992.
XX
PF 09-MAR-1992; 92WO-US01906.
XX
PR 07-MAR-1991; 91US-0666056.
PR 11-JUN-1991; 91US-0713967.
PR 06-MAR-1992; 92US-0847951.
XX
PA (VIRO-) VIROGENETICS CORP.
XX
PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP;
PI Limbach KJ, Norton EK, Paoletti E, Perkus ME, Pincus SE;
PI Riviere M, Tartaglia J, Taylor J;
XX
DR WPI; 1992-331718/40.
DR N-PSDB; AAQ29106.
XX
PT Vaccine comprises recombinant, attenuated pox-virus - use for
PT vaccinating against viral infections such as rabies, hepatitis B,
PT HIV, HSV, EBV, CMV, mumps etc.
XX
PS Disclosure; Fig 15; 456pp; English.
XX
CC The sequence given is encoded by an expression cassette which
CC comprises the hepatitis B virus (HBV) L protein (large pre-S antigen,
CC lpsAg) gene which is precisely linked to the Amsacta moorei
CC entomopoxvirus (AmEPV) 42 kD promoter. This DNA sequence was used in
CC the construction of a NYVAC recombinant expressing the HBV gene.
CC Other HBV genes were also used in the construction. These were HBV M

CC protein (small pre-S antigen, spsAg) and a fusion protein composed of
CC the entire pre-S region (S12/core, S1 + S2). Each of these gene
CC sequences were inserted individually into three different sites of
CC NYVAC separated by from each other by large regions of vaccinia DNA
CC containing essential genes. NYVAC is a Copenhagen vaccine strain of
CC vaccinia virus which has been modified by deletion of six non-essential
CC regions of the genome encoding known or potential virulence factors.
CC The deletion loci were engineered as recipient loci for the insertion
CC of foreign genes. The spacing of the three inserted sequences ensured
CC that any recombination that did occur would lead to disruption of the
CC vaccinia genome and would cause unviable vaccinia virus. See also
CC AAQ35501-864.
XX
SQ Sequence 389 AA;

Query Match 96.9%; Score 1870; DB 13; Length 389;
Best Local Similarity 98.0%; Pred. No. 4.2e-138;
Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:||||||||||||||:||||||||||||||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQTGRQPTPLSPPLRNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||||||||||||| ||||||| ||:||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGAVNPVPTTASPLSSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 MTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||||||||||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 389
```

RESULT 5
AAG62931
ID AAG62931 standard; Protein; 389 AA.
XX
AC AAG62931;
XX
DT 17-SEP-2001 (first entry)
XX
DE Amino acid sequence of a mutated HBV Pre-S protein.
XX
KW S gene; HBsm; mHVB; HBV; pre-S gene; C gene; P gene; X gene;
KW surface antigen; envelope protein; HBsAg; HBeAg; HBcAg; DNA polymerase;
KW reverse transcriptase; HBxAg; immune response; viral protein expression;
KW HBV infection.
XX
OS Hepatitis B virus.
XX
PN WO200149828-A1.
XX
PD 12-JUL-2001.
XX
PF 05-JAN-2001; 2001WO-FR00038.
XX
PR 06-JAN-2000; 2000FR-0000129.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (INMR) BIOMERIEUX SA.
XX
PI Trepo C, Mandrand B, Kay A, Chemin I, Komurian-Pradel F;
XX
DR WPI; 2001-441870/47.
XX
PT New mutant of hepatitis B virus, useful for diagnosis, prevention and
PT treatment of infection, also related nucleic acids, proteins and
PT antibodies -
XX
PS Claim 36; Page 92-94; 102pp; French.
XX
CC The present sequence represents a mutated Pre-S gene. It is part of the
CC mutant hepatitis B virus (mHVB) of the invention, which has a circular,
CC partially double-stranded genome comprising pre-S, S, C, P and X genes,
CC encoding, respectively, surface antigens, envelope protein (HBsAg),
CC proteins HBeAg and HBcAg, DNA polymerase/reverse transcriptase, and
CC HBxAg. The mutant HBV of the invention are used for induction of
CC specific humoral or cellular immune responses, or antisense inhibition
CC of viral protein expression. Proteins and nucleic acids derived from
CC mHVB (also related antibodies, vectors, transformed cells and
CC oligonucleotides) are useful for diagnosis (in standard amplification,
CC hybridization or immune-complex forming assays) and for treatment of,
CC or vaccination against, HBV infection.
XX
SQ Sequence 389 AA;

Query Match 96.8%; Score 1868; DB 22; Length 389;
Best Local Similarity 98.2%; Pred. No. 6.1e-138;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          ||||||||||||||||||||||||||||:|||||||||||||||||||||||| |||||
Db      48 GAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLGNTHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||||||||||||| |||||||||||||||||||| ||||||||||||||||||
Db     108 AMQWNSTTFHQTLQDPRVRGLYLPAGGSSSGTVNPVPTTVSHISSIFSRIGDPALNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          ||||||||||||||||||||||:|| ||||||||||||||||
Db     348 GLSPTVWLSVIWMMWYWGPSLYNILRPFLPLLPIFFCLWVYI 389
```

RESULT 6
AAU14097
ID AAU14097 standard; peptide; 393 AA.
XX
AC AAU14097;
XX
DT 21-NOV-2001 (first entry)
XX
DE Hepatitis B virus major surface antigen precursor S.
XX
KW Anti-retroviral; DP178-like; DP107-like; antifusogenic; antiviral;
KW major surface antigen precursor S; HIV transmission.
XX
OS Hepatitis B virus subtype AYW.
XX
PN WO200151673-A2.
XX
PD 19-JUL-2001.
XX
PF 05-JUL-2000; 2000WO-US35727.

XX
PR 09-JUL-1999; 99US-0350841.
XX
PA (TRIM-) TRIMERIS INC.
XX
PI Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
XX
DR WPI; 2001-442157/47.
XX
PT Identifying a compound that inhibits the formation of or disrupts a
PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral
PT or intracellular modulatory activity, by detecting the formation of a
PT DP107/DP178 complex -
XX
PS Disclosure; Fig 35; 259pp; English.
XX
CC The present invention relates to peptides which exhibit anti-retroviral
CC activity. The peptides of the invention (AAU12559-AAU14009) comprise
CC DP178-like and DP107-like peptides. The DP178 peptide corresponds
CC to amino acids 639-673 of the transmembrane protein gp41 from human
CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide
CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention
CC also relates to a method of identifying compounds that inhibit the
CC formation of or disrupts a DP107/DP178 complex. The method comprises
CC detecting the formation of a DP107/DP178 complex, both in the presence
CC or absence of a test compound, in a reaction mixture containing DP107
CC and DP178 peptides. The method is useful for identifying compounds,
CC including small molecule compounds, which may themselves exhibit
CC antifusogenic, antiviral or intracellular modulatory activity. The
CC DP178-like/DP107-like peptides are useful to inhibit human and non-human
CC retroviral, particularly HIV, transmission to uninfected cells. The
CC present sequence represents a peptide sequence from Hepatitis B virus
CC subtype AYW major surface antigen precursor S.
XX
SQ Sequence 393 AA;

Query Match 93.5%; Score 1803.5; DB 22; Length 393;
Best Local Similarity 96.5%; Pred. No. 6.7e-133;
Matches 333; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

```
Qy       6 GAGAFGLGFTPP-HGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNT-H 63
           |||||||||||| ||||||||||||||:||||||||||||||||||||||||||||||
Db      49 GAGAFGLGFTPPIIGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLRNTII 108

Qy      64 PQAMQWNSTTF-HQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNM 122
           |||||||||||  |||||||||||||||||||||||| || ||:|||||||||||||||
Db     109 PQAMQWNSTTFIIQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSSIFSRIGDPALNM 168

Qy     123 ENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHS 182
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     169 ENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHS 228

Qy     183 PTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPC 242
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 PTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPC 288

Qy     243 RTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQ 302
           ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 RTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQ 348

Qy     303 WFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||||||||||||||||||||||||||
Db     349 WFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 393
```

RESULT 7
AAG66932
ID AAG66932 standard; Protein; 399 AA.
XX
AC AAG66932;
XX
DT 19-OCT-2001 (first entry)
XX
DE HBV genotype E preS1/preS2/HBsAg polypeptide.
XX
KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW HBsAg; antiviral; vaccine; genotype G; genotyping; HBcAg; HBeAg.
XX
OS Hepatitis B virus.
XX
PN WO200140279-A2.
XX
PD 07-JUN-2001.
XX
PF 20-NOV-2000; 2000WO-EP11526.
XX
PR 03-DEC-1999; 99EP-0870252.
PR 07-DEC-1999; 99US-0169287.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy -
XX
PS Example 3; Fig 6; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype
CC G. This genotype was found with a high prevalence in patients
CC chronically infected with HBV and residing in Europe and the USA. The
CC invention relates to a fully defined sequence of 3248 nucleotides as
CC given in specification, a sequence with 92% identity to the given
CC sequence, or sequence that is degenerate to the mentioned sequences.
CC These polynucleotides are useful for HBV genotyping. The proteins
CC encoded by the polynucleotides are useful for detecting antibodies in
CC a biological sample. Ligands that bind to the proteins and antibodies
CC directed against the proteins are useful for detecting the proteins
CC and for detecting HBcAg and HBeAg (precore precursor proteins). They
CC are also useful for preparing a vaccine or medicament for treating
CC HBV infections. The present sequence is provided in an amino acid
CC sequence alignment of the preS1, preS2 and HBsAg open reading frame
CC of the different HBV genotypes.
XX
SQ Sequence 399 AA;

Query Match 92.4%; Score 1783; DB 22; Length 399;
Best Local Similarity 93.0%; Pred. No. 2.7e-131;
Matches 318; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           | |||| |||||||||||||||||:|:||||:||||||||||||||::|||||:||||
Db      58 GVGAFGPGFTPPHGGLLGWSPQAQGMLKTLPADPPPASTNRQSGRQPTPITPPLRDTHPQ 117

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||| |||||||||||||||||||||||| | |||||||||||| |||:|
Db     118 AMQWNSTTFHQALQDPRVRGLYFPAGGSSSGTVNPVPTTASLISSIPSRIGDPAPNMESI 177

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||:|||||||||||||||||  ||||||||||||||||||||
Db     178 TSGFLGPLLVLQAGFFLLTKILTIPQSLDSWWTSLNFLGGAPVCLGQNSQSPTSNHSPTS 237

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     238 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 297

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | |||||:||||||:||||||||||||||||||||||||||||||||||||||||||||
Db     298 MTLAQGTSMFPSCCCSKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFA 357

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
```

```
              ||||||||||||||||||||| |||||:|||||||||||||
Db     358 GLSPTVWLSVIWMMWYWGPSLYDILSPFIPLLPIFFCLWVYI 399

RESULT 8
AAR93801
ID   AAR93801 standard; Protein; 400 AA.
XX
AC   AAR93801;
XX
DT   22-OCT-1996  (first entry)
XX
DE   Hepatitis B virus pre S1/pre S2/S protein.
XX
KW   Hepatitis C virus; HCV; hepatitis B virus; HBV; fusion protein;
KW   immunisation; vaccine; infection; core protein; adr-1.
XX
OS   Hepatitis B virus.
XX
PN   WO9610997-A1.
XX
PD   18-APR-1996.
XX
PF   05-OCT-1995;   95WO-US13552.
XX
PR   06-JUN-1995;   95US-0467859.
PR   05-OCT-1994;   94US-0318248.
XX
PA   (APOL-) APOLLON INC.
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Coney LR,  Pachuk CJ,  Tokushige K,  Wakita T,  Wands J;
PI   Zurawski VR;
XX
DR   WPI; 1996-209642/21.
DR   N-PSDB; AAT36611.
XX
PT   Nucleic acid encoding hepatitis B and C virus fusion proteins - or
PT   incomplete hepatitis C virus genome, are useful in vaccines for
PT   prevention or treatment of HBV and HCV infections
XX
PS   Example 1; Page 32-34; 53pp; English.
XX
CC   A new fusion protein comprises a hepatitis B virus (HBV) S gene
CC   protein (AAT36611) coupled to amino acids 1-69, 1-70 or 1-154 of the
CC   hepatitis C virus (HCV) core protein (AAT28348).
CC   The nucleic acid has the coding sequence linked to a CMV promoter,
CC   RSV enhancer, polyadenylation sequence and opt. the 5'-UTR of HCV.
XX
SQ   Sequence   400 AA;

  Query Match             91.0%;  Score 1756;  DB 17;  Length 400;
  Best Local Similarity   91.2%;  Pred. No. 3.5e-129;
  Matches  312;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||| || |||||||||||||||||||| |:|| ||| ||||||||||||:|||||::|||
Db      59 GAGDFGPGFTPPHGGLLGWSPQAQGILTTVPAAPPPVSTNRQSGRQPTPISPPLRDSHPQ 118

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||| | ||||||||||||||||||||||||||||||||||||| |||| |||:
Db     119 AMQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRTGDPAPNMEST 178

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||    | ||||||||||||||
Db     179 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 238

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| ||||||||||||||||||||||||||||||||||||||||||:||:||||||||:||
Db     239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 298

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           | |||||||:||||||||||||||||||||||||| :||||||| |||||||||||||||
Db     299 TIPAQGTSMFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFA 358

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           |||||||||||||||||||||:||||||||||||||||||||
Db     359 GLSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 400

RESULT 9
AAP60794
ID   AAP60794 standard; Protein; 389 AA.
XX
AC   AAP60794;
XX
DT   30-OCT-1991  (first entry)
XX
DE   Adr-type hepatitis B virus SA, intermediate SA and large SA.
XX
KW   Surface antigen; vaccine; huGK-14.
XX
OS   adr-type Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..389
FT                   /label= 43K-47K large surface antigen
FT   Protein         109..389
FT                   /label= 31K-36K intermediate surface antigen
FT   Protein         164..389
FT                   /label= 22K-28K surface antigen
XX
PN   WO8603975-A.
XX
PD   17-JUL-1986.
XX
PF   28-DEC-1985;   85WO-JP00733.
XX
PR   28-DEC-1984;   84JP-0274592.
XX
PA   (NICA-) JAPAN FOUND CANCER.
PA   (MEIP ) MEIJI MILK PROD CO LTD.
PA   (GANK-) GAN KENKYU KAI ZH.
XX
PI   KOIKE K,  SUGANO H,  ODA M,  KATAYANAGI S;
XX
DR   WPI; 1986-196794/30.
DR   N-PSDB; AAN60714.
XX
PT   ADR-type hepatitis-B surface antigen for vaccine prodn. - obtd. by
PT   culture of human hepatoma cell strain hugk-14 avoiding use of human blood
XX
PS   Disclosure; Fig 5; 30pp; Japanese.
XX
CC   The sequence allows for the preparation of adr-type hepatitis B surface
CC   antigens for vaccine, without the use of human blood. The proteins may be
CC   expressed from human hepatoma cell strain huGK-14.
XX
SQ   Sequence   389 AA;

  Query Match             90.8%;  Score 1752;  DB 7;  Length 389;
  Best Local Similarity   90.9%;  Pred. No. 7e-129;
  Matches  311;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           ||| || |||||||||||||||||||| |:|| ||| ||||||||||||:|||||::|||
Db      48 GAGDFGPGFTPPHGGLLGWSPQAQGILTTVPAAPPPVSTNRQSGRQPTPISPPLRDSHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           |||||||||| | ||||||||||||||||||||||||||||||||||||| |||| |||:
Db     108 AMQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRTGDPAPNMEST 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||||    | ||||||||||||||
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 227
```

```
Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||| |||||||||||||||||||||||||||||||||||||||||:||:|||||||||:||
Db    228 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 287

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | |||||::|||||||||||||||||||||||| :|||||||| ||||||||||||||||
Db    288 TIPAQGTSIFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFA 347

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||:||||||||||||||||||||
Db    348 GLSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 389


RESULT 10
AAG66930
ID   AAG66930 standard; Protein; 400 AA.
XX
AC   AAG66930;
XX
DT   19-OCT-2001  (first entry)
XX
DE   HBV genotype C preS1/preS2/HBsAg polypeptide.
XX
KW   Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW   HBsAg; antiviral; vaccine; genotype G; genotyping; HBcAg; HBeAg.
XX
OS   Hepatitis B virus.
XX
PN   WO200140279-A2.
XX
PD   07-JUN-2001.
XX
PF   20-NOV-2000; 2000WO-EP11526.
XX
PR   03-DEC-1999;   99EP-0870252.
PR   07-DEC-1999;   99US-0169287.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Van Geyt C,  De Gendt S;
XX
DR   WPI; 2001-374785/39.
XX
PT   Novel isolated and/or purified hepatitis B virus polypeptide and
PT   polynucleotide sequences that are phylogenetically different from HBV
PT   genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT   therapy  -
XX
PS   Example 3; Fig 6; 94pp; English.
XX
CC   The invention relates to the complete nucleic acid sequence of a new
CC   human hepatitis B virus (HBV) genotype, provisionally named genotype
CC   G. This genotype was found with a high prevalence in patients
CC   chronically infected with HBV and residing in Europe and the USA. The
CC   invention relates to a fully defined sequence of 3248 nucleotides as
CC   given in specification, a sequence with 92% identity to the given
CC   sequence, or sequence that is degenerate to the mentioned sequences.
CC   These polynucleotides are useful for HBV genotyping. The proteins
CC   encoded by the polynucleotides are useful for detecting antibodies in
CC   a biological sample. Ligands that bind to the proteins and antibodies
CC   directed against the proteins are useful for detecting the proteins
CC   and for detecting HBcAg and HBeAg (precore precursor proteins). They
CC   are also useful for preparing a vaccine or medicament for treating
CC   HBV infections. The present sequence is provided in an amino acid
CC   sequence alignment of the preS1, preS2 and HBsAg open reading frame
CC   of the different HBV genotypes.
XX
SQ   Sequence   400 AA;

 Query Match           90.5%;  Score 1745;  DB 22;  Length 400;
 Best Local Similarity  90.9%;  Pred. No. 2.6e-128;
 Matches  311;  Conservative   11;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
          |||||| |||||||||||||||||:| |:|  |||||||||||||||||:||||::|||
Db     59 GAGAFGPGFTPPHGGLLGWSPQAQGVLTTVPVAPPPASTNRQSGRQPTPISPPLRDSHPQ 118

Qy     66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
          |||||||||| | ||||||||||||||||||||||||||| |||||  || |||| ||||
Db    119 AMQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTASPISSISSRTGDPAPNMENT 178

Qy    126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
          |||||||||||||||||||||||||||||||||||||||||||||    | ||||||||||||
Db    179 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 238

Qy    186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
          ||| |||||||||||||||||||||||||||||||||||||||||||:||:|||||||:||
Db    239 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 298

Qy    246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
          | |||||||:|||||||||||||||||||||||| :|||| || ||||||||||||||||
Db    299 TIPAQGTSMFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEGASVRFSWLSLLVPFVQWFV 358

Qy    306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
          |||||||||||||||||||||:||||||||||||||||||||
Db    359 GLSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 400


RESULT 11
AAG66919
ID   AAG66919 standard; Protein; 399 AA.
XX
AC   AAG66919;
XX
DT   19-OCT-2001  (first entry)
XX
DE   HBV HBPol protein.
XX
KW   Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW   HBsAg; antiviral; vaccine; genotype G; genotyping; HBcAg; HBeAg.
XX
OS   Hepatitis B virus.
XX
PN   WO200140279-A2.
XX
PD   07-JUN-2001.
XX
PF   20-NOV-2000; 2000WO-EP11526.
XX
PR   03-DEC-1999;   99EP-0870252.
PR   07-DEC-1999;   99US-0169287.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Van Geyt C,  De Gendt S;
XX
DR   WPI; 2001-374785/39.
DR   N-PSDB; AAH77562.
XX
PT   Novel isolated and/or purified hepatitis B virus polypeptide and
PT   polynucleotide sequences that are phylogenetically different from HBV
PT   genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT   therapy  -
XX
PS   Claim 16; Fig 1; 94pp; English.
XX
CC   The invention relates to the complete nucleic acid sequence of a new
CC   human hepatitis B virus (HBV) genotype, provisionally named genotype
CC   G. This genotype was found with a high prevalence in patients
CC   chronically infected with HBV and residing in Europe and the USA. The
CC   invention relates to a fully defined sequence of 3248 nucleotides as
CC   given in specification, a sequence with 92% identity to the given
CC   sequence, or sequence that is degenerate to the mentioned sequences.
CC   These polynucleotides are useful for HBV genotyping. The proteins
CC   encoded by the polynucleotides are useful for detecting antibodies in
CC   a biological sample. Ligands that bind to the proteins and antibodies
```

CC directed against the proteins are useful for detecting the proteins
CC and for detecting HBcAg and HBeAg (precore precursor proteins). They
CC are also useful for preparing a vaccine or medicament for treating
CC HBV infections. The present sequence is encoded by the genome of
CC HBV genotype G strain FR1.
XX
SQ Sequence 399 AA;

Query Match 90.4%; Score 1743; DB 22; Length 399;
Best Local Similarity 90.4%; Pred. No. 3.7e-128;
Matches 309; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
Db      58 GVGAYGPGFTPPHGGLLGWSPQSQGTLTTLPADPPPASTNRQSGRQPTPISPPLRDSHPQ 117

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
Db     118 AMQWNSTAFHQALQNPKVRGLYFPAGGSSSGIVNPVPTIASHISSIFSRIGDPAPNMENI 177

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
Db     178 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGVPVCPGLNSQSPTSNHSPIS 237

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
Db     238 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCKTC 297

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
Db     298 TTPAQGNSMYPSCCCTKPSDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFV 357

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
Db     358 GLSPTVWLSAIWMMWYWGPNLYNILSPFIPLLPIFFCLWVYI 399
```

RESULT 12
AAP80416
ID AAP80416 standard; protein; 383 AA.
XX
AC AAP80416;
XX
DT 12-SEP-1990 (first entry)
XX
DE Sequence of adr type hepatitis B virus (HBV) surface antigen
DE (HBsAg) L protein (M protein, S protein).
XX
KW Egg white lysozyme; hepatitis B virus (HBV) surface antigen(HBsAg);
KW Saccharomyces cerevisiae AH22R/pGLD LP39-RcT; L protein; M protein;
KW Saccharomyces cerevisiae L11P39-RcT; Saccharomyces cerevisiae LP31-RcT;
KW S protein.
XX
OS Hepatitis B virus.
XX
PN EP288198-A.
XX
PD 26-OCT-1988.
XX
PF 13-APR-1988; 88EP-0303297.
XX
PR 12-OCT-1987; 87JP-0256885.
XX
PA (TAKE) TAKEDA CHEMICAL IND KK.
XX
PI Fujisawa Y, Imai S, Miyazaki T;
XX
DR WPI; 1988-301233/43.
DR N-PSDB; AAN80973.
XX
PT Peptide(s) having hepatitis B surface antigenicity -
PT used as vaccine for prevention of hepatitis B virus infection
PT and in diagnostic kits
XX
PS Example; Fig 5; 31pp; English.
XX
CC HBsAg L protein (P39) is an env protein. When L protein genes are
CC introduced into animal cells (CHO cells) they produce HBsAg particles
CC containing M and S proteins. The advantage is that peptides having
CC HBsAg antigenicity are excreted outside the cells and easily purified.
CC Claimed is a eukaryotic cell transformed with the rDNA. Pref.
CC the cell is a yeast cell, esp. S. cerevisiae AH22R-/pGLD LP39-
CC RcT, L11P39-RcT or LP31-RcT. Also claimed is recombinant DNA, which
CC is a DNA coding for a signal peptide which functions in a eukaryotic
CC cell, and is bound to the 5'-terminal of a DNA coding for a peptide
CC having HBsAg activity. The signal peptide may be a signal peptide of
CC egg white lysozme. When trypsin-like protease-producing yeast is
CC utilized as the host, L protein and M protein may possibly be degraded
CC by the protease; therefore it is desirable that the genes are altered
CC so that the 48th arginine residue from the N-terminal of M protein or a
CC peptide containing the residue (preferably the 44-49 peptide) may be
CC deleted. Preferable genes include L protein gene (altered) coding for
CC the amino acid sequence 1-383 shown in AAN80973.
XX
SQ Sequence 383 AA;

Query Match 90.0%; Score 1736; DB 9; Length 383;
Best Local Similarity 90.6%; Pred. No. 1.2e-127;
Matches 310; Conservative 10; Mismatches 16; Indels 6; Gaps 1;

```
Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
Db      48 GAGAFGPGFTPPHGGLLGWSPQAQGILTTVPAAPPPASTNRQSGRQPTPISPPLRDSHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
Db     108 AMQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTASPIS------GDPAPNMENT 161

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
Db     162 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 221

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
Db     222 CPPICPGYRWMCLRRFIIFLPILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 281

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
Db     282 TIPAQGTSMFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFV 341

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
Db     342 GLSPTVWLSVIWMMWYWGPSLYNILSPFLPLLPIFFCLWVYI 383
```

RESULT 13
AAR55287
ID AAR55287 standard; Protein; 389 AA.
XX
AC AAR55287;
XX
DT 02-JAN-1995 (first entry)
XX
DE Deduced sequence of HBV surface antigen LS.
XX
KW Surface antigen; recombinant replicable vaccinia virus; hepatitis;
KW prevention; therapy; epitope; hepatitis B virus.
XX
OS Hepatitis B virus.
XX
PN WO9412617-A.
XX
PD 09-JUN-1994.
XX
PF 24-NOV-1993; 93WO-US11474.
XX

```
PR   25-NOV-1992;   92US-0982211.
XX
PA   (ITBI-) INT BIOTECHNOLOGY LAB INC.
XX
PI   Bernstine EG,  Lewis T,  Okeefe RW,  Souw PTS;
XX
DR   WPI; 1994-200247/24.
DR   N-PSDB; AAQ67970.
XX
PT   Prevention and treatment of hepatitis - using recombinant
PT   replicable vaccinia viruses contg. hepatitis B virus surface and
PT   core antigen nucleotide sequences
XX
PS   Example; Fig 22; 252pp; English.
XX
CC   The HBV genome (subtype adw) is contained in plasmid pAM6. The
CC   3' end of the ORF conding for the three surface Ags was manipulated
CC   to generate pRO-02 which contains an intact ORF capable of encoding
CC   LS, MS and S Ags. The DNA sequences encoding S, MS and LS were sub-
CC   cloned and sequenced.  The sequences of the LS are AAQ67970/R55287;
CC   the sequences of MS are AAQ67961/R55281 and The sequences of S are
CC   AAQ67962/R55282. Changes in nucleotide and deduced AA sequences
CC   from the published sequence are believed to result from variability
CC   within the adw subtype.
XX
SQ   Sequence   389 AA;

  Query Match             89.5%;  Score 1727;  DB 15;  Length 389;
  Best Local Similarity   89.2%;  Pred. No. 6.3e-127;
  Matches  305;  Conservative   14;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
Db      48 GVGAFGPGFTPPHGGVLGWSPQAQGILTTVSTIPPPASTNRQSGRQPTPISPPLRDSHPQ 107

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
Db     108 AMQWNSTAFHQALQDPRVRGLYFPAGGSSSGTVNPAPNIASHISSISARTGDPVTNMENI 167

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
Db     168 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGSPVCLGQNSQSPTSNHSPTS 227

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
Db     228 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSTTTSTGPCKTC 287

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
Db     288 TTPAQGNSMFPSCCCTKPTDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFV 347

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
Db     348 GLSPTVWLSAIWMMWYWGPSLYSIVSPFIPLLPIFFCLWVYI 389


RESULT 14
AAY54045
ID   AAY54045 standard; Protein; 400 AA.
XX
AC   AAY54045;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Amino acid seqeunce of a HBV large surface antigen protein.
XX
KW   HBV; HBV surface antigen-'S'-133 Oon strain (Met to Thr);
KW   DNA polymerase; large surface antigen; core protein;
KW   transactivating X protein; hepatitis vaccine; HBV infection;
KW   hepatocellular carcinoma.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 14
FT                   /note= "encoded by AGC"
FT   Misc-difference 174..400
FT                   /note= " these residues are specifically claimed in
FT                     claim 29"
FT   Misc-difference 298..320
FT                   /note= " these residues are specifically claimed in
FT                     claim 23"
XX
PN   WO9966048-A1.
XX
PD   23-DEC-1999.
XX
PF   19-JUN-1998;   98WO-SG00046.
XX
PR   19-JUN-1998;   98WO-SG00046.
XX
PA   (GOVE-) GOVERNMENT REPUBLIC SINGAPORE.
XX
PI   Oon CJ,  Lim GK,  Zhao Y,  Chen WN;
XX
DR   WPI; 2000-106104/09.
DR   N-PSDB; AAZ37088.
XX
PT   New isolated hepatitis B virus strain, useful for, e.g. treatment of
PT   hepatitis infection -
XX
PS   Claim 23; Page 39-40; 68pp; English.
XX
CC   The present sequence is encoded by the genome of an isolated strain of
CC   Hepatitis B virus designated human Hepatitis B virus (HBV) surface
CC   antigen-'S'-133 Oon strain (Met to Thr). The viral genome is deposited
CC   as ECCC accession numbers P97121501, P97121502 and P97121503. The
CC   nucleotide sequence was isolated from hepatocellular carcinoma (HCC).
CC   The nucleotide sequence encodes four overlapping proteins, which are
CC   a DNA polymerase, a large surface antigen, a core protein, and a
CC   transactivating X protein. The large surface antigen differs from the
CC   wild type sequence in that it contains a Thr at position 133 of the
CC   wild type sequence instead of a Met. The proteins are used to produce
CC   antibodies. The proteins, polynucleotide and antibodies can be used for
CC   detecting the novel HBV strain. The HBV polypeptides can also be used
CC   in hepatitis vaccines. The HBV novel strain polypeptides can be used
CC   to identify compounds for treating or preventing HBV infection or
CC   hepatocellular carcinoma.
XX
SQ   Sequence   400 AA;

  Query Match             89.2%;  Score 1721;  DB 21;  Length 400;
  Best Local Similarity   90.1%;  Pred. No. 1.9e-126;
  Matches  308;  Conservative   10;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
Db      59 GVGAFGPGFTPPHGGLLGGSPQAQGILTTVPAAPPPASTNRQSGRQPTPISPPLRDSHPQ 118

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
Db     119 ATQWNSTTFHQALLDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRTGDPAPNMENT 178

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
Db     179 TSGFLGPLLVLQAGFFSLTRILTIPQSLDSWWTSLNFLGGAPTCPGQNSQSPTSNHSPTS 238

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
Db     239 CPPICPGYRWNCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLLPGTSTTSTGPCKTC 298

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
Db     299 TTPAQGNSTFPSCCCTKPSDGNCTCIPIPSSWAFARFLWEWASVRFSWLSLLVPFVQWFV 358

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
```

```
                  |||||||||||||||||| |||:||||||||||||||||||||
Db    359 GLSPTVWLSVIWMMWYWGRSLYNILSPFLPLLPIFFCLWVYI 400


RESULT 15
AAR10851
ID   AAR10851 standard; Protein; 388 AA.
XX
AC   AAR10851;
XX
DT   29-APR-1991  (first entry)
XX
DE   Modified hepatitis B virus large protein (I).
XX
KW   Hepatitis B virus; large surface protein; L protein; myristilation;
KW   vaccines.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   Region          1..107
FT                   /label= pre-S1
FT   Region          108..162
FT                   /label= pre-S2
FT   Region          163..388
FT                   /label= S-region
XX
PN   EP414374-A.
XX
PD   27-FEB-1991.
XX
PF   19-JUL-1990;   90EP-0307900.
XX
PR   03-AUG-1989;   89US-0389184.
PR   25-JUL-1989;   89US-0385342.
XX
PA   (SMIK ) SMITHKLINE BIOLOGIC.
XX
PI   Comberbach M,  Harford N,  Cabezon T,  Rutgers A,  Voet P;
PI   Jacobs E,  Hollenberg CP,  Janowicz ZA,  Merckelberg  AJ;
XX
DR   WPI; 1991-059585/09.
DR   N-PSDB; AAQ10678.
XX
PT   Modified hepatitis B virus large surface protein - has amino acid
PT   sequence encoding L protein used in vaccine for treating or
PT   preventing hepatitis B without effects
XX
PS   Disclosure; Page 4-7; 79pp; English.
XX
CC   The modified L protein has the Gly 13 amino acid in the
CC   preS1 region of the wild-type sequence deleted.  The deletion
CC   results in the synthesis of a non-myristilated L protein.
CC   (The first amino of the ad sequence are not given).
CC   The modified product may be used alone or in a composite particle
CC   comprising at least two polypeptides corresp. to all or part of a
CC   protein having the biological activity of one of the hepatitis
CC   B surface antigens.  The composite particles are useful for
CC   preparing improved hepatitis B vaccines or for treating of HBV.
CC   See also AAQ10677-79.
XX
SQ   Sequence   388 AA;

  Query Match             88.8%;  Score 1713;  DB 12;  Length 388;
  Best Local Similarity   88.3%;  Pred. No. 7.8e-126;
  Matches  302;  Conservative   15;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       6 GAGAFGLGFTPPHGGLLGWSPQAQGILETLPANPPPASTNRQSGRQPTPLSPPLRNTHPQ 65
           | |||| | ||||||:|||||||||||| |:   |||||||||||||||:|||||::|||
Db      47 GVGAFGPGLTPPHGGILGWSPQAQGILTTVSTIPPPASTNRQSGRQPTPISPPLRDSHPQ 106

Qy      66 AMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALNMENI 125
           ||||||| ||| ||||||||||||||||||||||| |   | |||  :| |||  |||||

Db     107 AMQWNSTAFHQALQDPRVRGLYFPAGGSSSGTVNPAPNIASHISSSSARTGDPVTNMENI 166

Qy     126 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTS 185
           ||||||||||||||||||||||||||||||||||||||: |||||||||||||||||
Db     167 TSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGSPVCLGQNSQSPTSNHSPTS 226

Qy     186 CPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTC 245
           ||| |||||||||||||||||||||||||||||||||||||||||||:||:||||:||
Db     227 CPPICPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSTTTNTGPCKTC 286

Qy     246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQWFV 305
           |||||| ||:||||||||:|||||||||||||| |:|||||| |||||||||||||||
Db     287 TTPAQGNSMFPSCCCTKPTDGNCTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFV 346

Qy     306 GLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI 347
           ||||||||| ||||||||||||||:|||:|||||||||||||
Db     347 GLSPTVWLSAIWMMWYWGPSLYSIVSPFIPLLPIFFCLWVYI 388


Search completed: February  3, 2003, 09:23:50
Job time : 57.6167 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:12 ; Search time 28.8415 Seconds
(without alignments)
1535.988 Million cell updates/sec

Title: US-09-890-752A-2
Perfect score: 1148
Sequence: 1 MPLSSIFSRIGDPTVQASKL.....RRRRSQSPRRRRSQSREPQC 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1048.5	91.3	212	12	O11884	O11884 hepatitis b
2	1040.5	90.6	212	12	Q89656	Q89656 hepatitis b
3	1038.5	90.5	212	12	Q68016	Q68016 hepatitis b
4	1037.5	90.4	212	12	Q67876	Q67876 hepatitis b
5	1037.5	90.4	212	12	Q68020	Q68020 hepatitis b
6	1037.5	90.4	212	12	Q89597	Q89597 hepatitis b
7	1035.5	90.2	212	12	Q68025	Q68025 hepatitis b
8	1035.5	90.2	212	12	Q68068	Q68068 hepatitis b
9	1034.5	90.1	212	12	Q67984	Q67984 hepatitis b
10	1034.5	90.1	212	12	Q68032	Q68032 hepatitis b
11	1034.5	90.1	212	12	Q68070	Q68070 hepatitis b
12	1032.5	89.9	212	12	Q91SY7	Q91sy7 hepatitis b
13	1032.5	89.9	212	12	Q917I7	Q917i7 hepatitis b
14	1032.5	89.9	212	12	Q910K2	Q910k2 hepatitis b
15	1031.5	89.9	212	12	Q67872	Q67872 hepatitis b
16	1031.5	89.9	212	12	Q68051	Q68051 hepatitis b
17	1030.5	89.8	212	12	Q91T01	Q91t01 hepatitis b
18	1030.5	89.8	212	12	Q91SX4	Q91sx4 hepatitis b
19	1030.5	89.8	212	12	Q67980	Q67980 hepatitis b
20	1030.5	89.8	212	12	Q68012	Q68012 hepatitis b
21	1030.5	89.8	212	12	Q68014	Q68014 hepatitis b
22	1029.5	89.7	212	12	Q91SZ5	Q91sz5 hepatitis b
23	1029.5	89.7	212	12	Q91SZ1	Q91sz1 hepatitis b
24	1029.5	89.7	212	12	Q91SY4	Q91sy4 hepatitis b
25	1029.5	89.7	212	12	Q910U3	Q910u3 hepatitis b
26	1029.5	89.7	212	12	Q910N8	Q910n8 hepatitis b
27	1029.5	89.7	212	12	Q68010	Q68010 hepatitis b
28	1029.5	89.7	212	12	Q68075	Q68075 hepatitis b
29	1028.5	89.6	212	12	Q96845	Q96845 hepatitis b
30	1028.5	89.6	212	12	Q9QMH8	Q9qmh8 hepatitis b
31	1028.5	89.6	212	12	Q917J5	Q917j5 hepatitis b
32	1028.5	89.6	212	12	Q68030	Q68030 hepatitis b
33	1028.5	89.6	212	12	Q89446	Q89446 hepatitis b
34	1027.5	89.5	212	12	Q91SX8	Q91sx8 hepatitis b
35	1027.5	89.5	212	12	Q910L4	Q910l4 hepatitis b
36	1026.5	89.4	212	12	Q91T03	Q91t03 hepatitis b
37	1026.5	89.4	212	12	Q91SZ6	Q91sz6 hepatitis b
38	1026.5	89.4	212	12	Q91SZ0	Q91sz0 hepatitis b
39	1026.5	89.4	212	12	Q91SY8	Q91sy8 hepatitis b
40	1026.5	89.4	212	12	Q91SY6	Q91sy6 hepatitis b
41	1026.5	89.4	212	12	Q91SY5	Q91sy5 hepatitis b
42	1026.5	89.4	212	12	Q91SY3	Q91sy3 hepatitis b
43	1026.5	89.4	212	12	Q917J7	Q917j7 hepatitis b
44	1026.5	89.4	212	12	Q910L8	Q910l8 hepatitis b
45	1025.5	89.3	212	12	Q91T04	Q91t04 hepatitis b

ALIGNMENTS

RESULT 1
O11884
ID O11884 PRELIMINARY; PRT; 212 AA.
AC O11884;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Core antigen.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Rao B.S., Casey J.L., Rinaudo J.S., Korba B.E.;
RT "Complete nucleotide sequence of a molecular clone of hepatitis B virus from the 2.2.15 cell line.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U95551; AAB59971.1; -.
DR InterPro; IPR002006; Hepatitis_core.
DR Pfam; PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24360 MW; D301689E8B05A1B8 CRC64;

Query Match 91.3%; Score 1048.5; DB 12; Length 212;
Best Local Similarity 97.5%; Pred. No. 2.5e-98;
Matches 198; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
```

```
                    ||||||||||||||||||||||
Db    190 PSPRRRRSQSPRRRRSQSREPQC 212


RESULT 2
Q89656
ID   Q89656      PRELIMINARY;      PRT;   212 AA.
AC   Q89656;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Core antigen (Precore protein).
GN   PREC.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AYW4;
RA   Pluciennicak A.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81012091; PubMed=399327;
RA   Galibert F., Mandart E., Fitoussi F., Tiollais P., Charnay P.;
RT   "Nucleotide sequence of the hepatitis B virus genome (subtype ayw)
RT   cloned in E. coli.";
RL   Nature 281:646-650(1979).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Borisova G.P., Pumpen P.P., Bychko V.V., Pushko P.M., Kalis Y.V.,
RA   Dishler A.V., Gren E.Y., Tsibinogin V.V., Kukain R.A.;
RL   Dokl. Biochem. 279:386-390(1985).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RA   Petzold D.R., Tautz B., Wolf F., Drescher J.;
RT   "Infection chains and evolution rates of Hepatitis B Virus in cardiac
RT   transplant recipients infected nosocommially.";
RL   J. Med. Virol. 0:0-0(1999).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RA   Petzold D.R.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RA   Kay A.C.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE FROM N.A.
RA   Jeantet D., Chemin I., Mandrand B., Zoulim F., Trepo C., Kay A.;
RT   "Characterization of two Hepatitis B virus populations isolated from a
RT   Hepatitis B surface antigen-negative patient.";
RL   Hepatology 0:0-0(0).
DR   EMBL; X85290; CAA59609.1; -.
DR   EMBL; X85300; CAA59631.1; -.
DR   EMBL; X85313; CAA59659.1; -.
DR   EMBL; X85283; CAA59593.1; -.
DR   EMBL; X85306; CAA59644.1; -.
DR   EMBL; Z35716; CAA84786.1; -.
DR   EMBL; X85312; CAA59657.1; -.
DR   EMBL; AJ131956; CAA10538.1; -.
DR   EMBL; AJ344117; CAC87011.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
SQ   SEQUENCE   212 AA;  24350 MW;  D3016A4E8B05A1B8 CRC64;

  Query Match             90.6%;  Score 1040.5;  DB 12;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 1.6e-97;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||      ||||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 3
Q68016
ID   Q68016      PRELIMINARY;      PRT;   212 AA.
AC   Q68016;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Pre-c/core protein.
GN   PRE-C/CORE.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X85288; CAA59604.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
SQ   SEQUENCE   212 AA;  24316 MW;  D1531ABF1B234208 CRC64;

  Query Match             90.5%;  Score 1038.5;  DB 12;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 2.6e-97;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRALLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||      ||||||||||||||||||||:
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQI 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||||||
Db     190 PSPRRRRSQSPRRRRSQSREPQC 212


RESULT 4
Q67876
ID   Q67876      PRELIMINARY;      PRT;   212 AA.
AC   Q67876;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Pre C/C ORF.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lai M.E., Mazzoleni A.P., Balestrieri A., Melis A., Porru A.;
RT   "Sequence analysis of HBV genomes isolated from patients with HBsAg
RT   negative chronic liver disease.";
RL   Submitted (MAR-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X65258; CAA46354.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
SQ   SEQUENCE   212 AA;  24348 MW;  6E91CC7D068EB573 CRC64;

  Query Match          90.4%;  Score 1037.5;  DB 12;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 3.3e-97;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVAFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 5
Q68020
ID   Q68020      PRELIMINARY;      PRT;   212 AA.
AC   Q68020;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Pre-c/core protein.
GN   PRE-C/CORE.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X85291; CAA59611.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
SQ   SEQUENCE   212 AA;  24336 MW;  63006A4E8B04A1B8 CRC64;

  Query Match          90.4%;  Score 1037.5;  DB 12;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 3.3e-97;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||:|||||||||||||      ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGDLMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 6
Q89597
ID   Q89597      PRELIMINARY;      PRT;   212 AA.
AC   Q89597;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HBCAG (PRE-CORE protein) (PRECORE/core protein).
OS   Hepatitis B virus (subtype ayw), and
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10418, 10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Hepatitis B virus (subtype ayw); STRAIN=SUB-TYPE AYW;
RX   MEDLINE=94079539; PubMed=8257295;
RA   Preisler-Adams S., Schlayer M.J., Peters T., Hettler F., Gerok W.,
RA   Rasenack J.;
RT   "Sequence analysis of hepatitis B virus DNA in immunologically
RT   negative infection.";
RL   Arch. Virol. 133:385-396(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Hepatitis B virus (subtype ayw); STRAIN=AYW;
RA   Karayiannis P.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Hepatitis B virus (subtype ayw); STRAIN=AYW;
RA   Karayiannis P.;
RL   Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Hepatitis B virus; STRAIN=I43-FAM G;
RA   Zampino R., Lobello S., Chiaramonte M., Venturi-Pasini C., Dumpis U.,
RA   Thursz M., Karayiannis P.;
RT   "Intra-familial transmission of Hepatitis B virus in Italy:
RT   Phylogenetic sequence analysis and amino acid variation of the core
RT   gene.";
RL   J. Hepatol. 0:0-0(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Hepatitis B virus; STRAIN=GAM1119F29, AND GAM1821F217;
RA   Dumpis U., Mendy M., Karayiannis P.;
RT   "Prevalence of HBV core promoter/precore/core mutations in Gambian
RT   chronic carriers.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X72702; CAA51257.1; -.
DR   EMBL; X80925; CAA56887.1; -.
DR   EMBL; AF419525; AAL15953.1; -.
DR   EMBL; AF350127; AAK57244.1; -.
DR   EMBL; AF350205; AAK57322.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
FT   CHAIN        30    212       CORE PROTEIN.
SQ   SEQUENCE   212 AA;  24336 MW;  1861B13E8B047AC8 CRC64;

  Query Match          90.4%;  Score 1037.5;  DB 12;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 3.3e-97;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189
```

```
Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          ||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 7
Q68025
ID Q68025 PRELIMINARY; PRT; 212 AA.
AC Q68025;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Pre-c/core protein.
GN PRE-C/CORE.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; X85293; CAA59616.1; -.
DR InterPro; IPR002006; Hepatitis_core.
DR Pfam; PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24363 MW; 6940D04E90A410C3 CRC64;

Query Match 90.2%; Score 1035.5; DB 12; Length 212;
Best Local Similarity 96.6%; Pred. No. 5.3e-97;
Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

```
Qy     13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDNASALYREALESP 74

Qy     73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          |||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db     75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy    133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          ||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 8
Q68068
ID Q68068 PRELIMINARY; PRT; 212 AA.
AC Q68068;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Pre-c/core protein.
GN PRE-C/CORE.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; X85315; CAA59664.1; -.
DR InterPro; IPR002006; Hepatitis_core.
DR Pfam; PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24366 MW; CEABDD4E8B11E5E8 CRC64;

Query Match 90.2%; Score 1035.5; DB 12; Length 212;
Best Local Similarity 96.6%; Pred. No. 5.3e-97;
Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

```
Qy     13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy     73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          |||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db     75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy    133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 LWFHICCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          ||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 9
Q67984
ID Q67984 PRELIMINARY; PRT; 212 AA.
AC Q67984;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Pre-c/core protein.
GN PRE-C/CORE.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; X85256; CAA59519.1; -.
DR InterPro; IPR002006; Hepatitis_core.
DR Pfam; PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24322 MW; 629E0B298B13D0D1 CRC64;

Query Match 90.1%; Score 1034.5; DB 12; Length 212;
Best Local Similarity 96.6%; Pred. No. 6.7e-97;
Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

```
Qy     13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLSTASALYREALESP 74

Qy     73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          |||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db     75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy    133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          ||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 10
Q68032
ID Q68032 PRELIMINARY; PRT; 212 AA.
AC Q68032;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Pre-c/core protein.
GN PRE-C/CORE.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.

```
RA   Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X85296; CAA59622.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
SQ   SEQUENCE   212 AA;  24308 MW;  E28972C78B1604FF CRC64;

  Query Match             90.1%;  Score 1034.5;  DB 12;  Length 212;
  Best Local Similarity   96.1%;  Pred. No. 6.7e-97;
  Matches  195;  Conservative    2;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||:||||||||||||||     ||||||||||||||||||:||
Db      75 EHCSPHHTALRQAILCWGDLMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFKQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 11
Q68070
ID   Q68070      PRELIMINARY;      PRT;   212 AA.
AC   Q68070;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Pre-c/core protein.
GN   PRE-C/CORE.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X85316; CAA59666.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
SQ   SEQUENCE   212 AA;  24376 MW;  D30173662996A1B8 CRC64;

  Query Match             90.1%;  Score 1034.5;  DB 12;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 6.7e-97;
  Matches  196;  Conservative    0;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||     ||||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILLTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 12
Q91SY7
ID   Q91SY7      PRELIMINARY;      PRT;   212 AA.
AC   Q91SY7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Precore/core protein.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GAM1211;
RA   Dumpis U., Mendy M., Karayiannis P.;
RT   "Prevalence of HBV core promoter/precore/core mutations in Gambian
RT   chronic carriers.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF350155; AAK57272.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
SQ   SEQUENCE   212 AA;  24363 MW;  27E96F6D069BC35A CRC64;

  Query Match             89.9%;  Score 1032.5;  DB 12;  Length 212;
  Best Local Similarity   96.1%;  Pred. No. 1.1e-96;
  Matches  195;  Conservative    1;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||     ||||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPENTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 13
Q917I7
ID   Q917I7      PRELIMINARY;      PRT;   212 AA.
AC   Q917I7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Pre-core protein.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=I44-FAM G;
RA   Zampino R., Lobello S., Chiaramonte M., Venturi-Pasini C., Dumpis U.,
RA   Thursz M., Karayiannis P.;
RT   "Intra-familial transmission of Hepatitis B virus in Italy:
RT   Phylogenetic sequence analysis and amino acid variation of the core
RT   gene.";
RL   J. Hepatol. 0:0-0(2002).
DR   EMBL; AF419526; AAL15954.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
FT   CHAIN        30    212       CORE PROTEIN.
SQ   SEQUENCE   212 AA;  24421 MW;  B76FFE275B047AD1 CRC64;

  Query Match             89.9%;  Score 1032.5;  DB 12;  Length 212;
  Best Local Similarity   96.1%;  Pred. No. 1.1e-96;
  Matches  195;  Conservative    1;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
```

```
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||| |||||||||||||||||||||||||||     |||||||||||||||||||
Db      75 EHCSPHHTRLRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 14
Q910K2
ID   Q910K2      PRELIMINARY;      PRT;   212 AA.
AC   Q910K2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Precore/core protein.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VARIOUS STRAINS;
RA   Dumpis U., Mendy M., Karayiannis P.;
RT   "Prevalence of HBV core promoter/precore/core mutations in Gambian
RT   chronic carriers.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF350218; AAK57335.1; -.
DR   EMBL; AF350099; AAK57216.1; -.
DR   EMBL; AF350100; AAK57217.1; -.
DR   EMBL; AF350102; AAK57219.1; -.
DR   EMBL; AF350103; AAK57220.1; -.
DR   EMBL; AF350106; AAK57223.1; -.
DR   EMBL; AF350108; AAK57225.1; -.
DR   EMBL; AF350116; AAK57233.1; -.
DR   EMBL; AF350118; AAK57235.1; -.
DR   EMBL; AF350119; AAK57236.1; -.
DR   EMBL; AF350120; AAK57237.1; -.
DR   EMBL; AF350121; AAK57238.1; -.
DR   EMBL; AF350122; AAK57239.1; -.
DR   EMBL; AF350124; AAK57241.1; -.
DR   EMBL; AF350125; AAK57242.1; -.
DR   EMBL; AF350126; AAK57243.1; -.
DR   EMBL; AF350128; AAK57245.1; -.
DR   EMBL; AF350130; AAK57247.1; -.
DR   EMBL; AF350132; AAK57249.1; -.
DR   EMBL; AF350133; AAK57250.1; -.
DR   EMBL; AF350138; AAK57255.1; -.
DR   EMBL; AF350140; AAK57257.1; -.
DR   EMBL; AF350141; AAK57258.1; -.
DR   EMBL; AF350143; AAK57260.1; -.
DR   EMBL; AF350147; AAK57264.1; -.
DR   EMBL; AF350148; AAK57265.1; -.
DR   EMBL; AF350149; AAK57266.1; -.
DR   EMBL; AF350150; AAK57267.1; -.
DR   EMBL; AF350151; AAK57268.1; -.
DR   EMBL; AF350152; AAK57269.1; -.
DR   EMBL; AF350156; AAK57273.1; -.
DR   EMBL; AF350157; AAK57274.1; -.
DR   EMBL; AF350162; AAK57279.1; -.
DR   EMBL; AF350163; AAK57280.1; -.
DR   EMBL; AF350165; AAK57282.1; -.
DR   EMBL; AF350169; AAK57286.1; -.
DR   EMBL; AF350170; AAK57287.1; -.
DR   EMBL; AF350171; AAK57288.1; -.
DR   EMBL; AF350172; AAK57289.1; -.
DR   EMBL; AF350174; AAK57291.1; -.
DR   EMBL; AF350178; AAK57295.1; -.
DR   EMBL; AF350181; AAK57298.1; -.
DR   EMBL; AF350185; AAK57302.1; -.
DR   EMBL; AF350186; AAK57303.1; -.
DR   EMBL; AF350189; AAK57306.1; -.
DR   EMBL; AF350190; AAK57307.1; -.
DR   EMBL; AF350192; AAK57309.1; -.
DR   EMBL; AF350193; AAK57310.1; -.
DR   EMBL; AF350194; AAK57311.1; -.
DR   EMBL; AF350195; AAK57312.1; -.
DR   EMBL; AF350196; AAK57313.1; -.
DR   EMBL; AF350197; AAK57314.1; -.
DR   EMBL; AF350198; AAK57315.1; -.
DR   EMBL; AF350199; AAK57316.1; -.
DR   EMBL; AF350200; AAK57317.1; -.
DR   EMBL; AF350201; AAK57318.1; -.
DR   EMBL; AF350210; AAK57327.1; -.
DR   EMBL; AF350212; AAK57329.1; -.
DR   EMBL; AF350213; AAK57330.1; -.
DR   EMBL; AF350214; AAK57331.1; -.
DR   EMBL; AF350216; AAK57333.1; -.
DR   EMBL; AF350217; AAK57334.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
SQ   SEQUENCE   212 AA;  24349 MW;  9A61B13E8B10C583 CRC64;

  Query Match             89.9%;  Score 1032.5;  DB 12;  Length 212;
  Best Local Similarity   96.1%;  Pred. No. 1.1e-96;
  Matches  195;  Conservative    1;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||     |||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPENTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 15
Q67872
ID   Q67872      PRELIMINARY;      PRT;   212 AA.
AC   Q67872;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Pre C/C ORF.
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lai M.E., Mazzoleni A.P., Balestrieri A., Melis A., Porru A.;
RT   "Sequence analysis of HBV genomes isolated from patients with HBsAg
RT   negative chronic liver disease.";
RL   Submitted (MAR-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X65257; CAA46350.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
SQ   SEQUENCE   212 AA;  24348 MW;  1861AB0DA21E49E1 CRC64;

  Query Match             89.9%;  Score 1031.5;  DB 12;  Length 212;
  Best Local Similarity   96.1%;  Pred. No. 1.3e-96;
  Matches  195;  Conservative    1;  Mismatches    2;  Indels    5;  Gaps    1;
```

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLIFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

Search completed: February 3, 2003, 09:25:44
Job time : 30.1748 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:16:26 ; Search time 7.86585 Seconds
(without alignments)
1133.686 Million cell updates/sec

Title: US-09-890-752A-2
Perfect score: 1148
Sequence: 1 MPLSSIFSRIGDPTVQASKL.....RRRRSQSPRRRRSQSREPQC 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	992.5	86.5	214	1	CORA_HPBV9	P17099 hepatitis b
2	988	86.1	211	1	CORA_HPBVA	P24023 hepatitis b
3	985.5	85.8	212	1	CORA_HPBVT	Q05495 hepatitis b
4	951.5	82.9	183	1	CORA_HPBVY	P03146 hepatitis b
5	932.5	81.2	183	1	CORA_HPBVZ	P03147 hepatitis b
6	919.5	80.1	183	1	CORA_HPBVJ	P17391 hepatitis b
7	919.5	80.1	183	1	CORA_HPBVO	P17392 hepatitis b
8	919.5	80.1	185	1	CORA_HPBVW	P03149 hepatitis b
9	918.5	80.0	183	1	CORA_HPBV4	P03150 hepatitis b
10	912.5	79.5	195	1	CORA_HPBVF	P29178 hepatitis b
11	911.5	79.4	185	1	CORA_HPBV2	P03148 hepatitis b
12	899.5	78.4	183	1	CORA_HPBVL	P12901 hepatitis b
13	757.5	66.0	217	1	CORA_HPBGS	P03153 ground squi
14	671	58.4	188	1	CORA_WHV1	P03152 woodchuck h
15	655.5	57.1	187	1	CORA_WHV8	P06433 woodchuck h
16	175	15.2	305	1	CORA_HPBHE	P13845 heron hepat
17	173	15.1	305	1	CORA_HPBDC	P30027 duck hepati
18	171	14.9	305	1	CORA_HPBDB	P17190 duck hepati
19	168	14.6	305	1	CORA_HPBDW	P17191 duck hepati
20	166	14.5	305	1	CORA_HPBDU	P03154 duck hepati
21	111	9.7	196	1	SFR2_CAEEL	Q09511 caenorhabdi
22	94.5	8.2	484	1	SFRB_HUMAN	Q05519 homo sapien
23	90.5	7.9	208	1	YSX2_CAEEL	Q10021 caenorhabdi
24	86	7.5	951	1	SFR8_HUMAN	Q12872 homo sapien
25	84	7.3	219	1	NOL3_HUMAN	O60936 homo sapien
26	84	7.3	498	1	VE2_HPV08	P06422 human papil
27	84	7.3	503	1	VE2_HPV21	P50767 human papil
28	80.5	7.0	329	1	SRA4_CAEEL	Q09206 caenorhabdi
29	80.5	7.0	3148	1	HD_FUGRU	P51112 fugu rubrip
30	80	7.0	832	1	ANR3_HUMAN	P57078 homo sapien
31	80	7.0	867	1	VL96_IRV1	P22856 tipula irid
32	79.5	6.9	303	1	SFR1_ARATH	O22315 arabidopsis
33	79.5	6.9	891	1	MAY1_SCHCO	P37936 schizophyll
34	79	6.9	483	1	VE2_HPV14	P36783 human papil
35	79	6.9	2799	1	EDD_HUMAN	O95071 homo sapien
36	78	6.8	164	1	SFR3_HUMAN	P23152 homo sapien
37	78	6.8	1042	1	SUWA_DROME	P12297 drosophila
38	78	6.8	2426	1	SON_HUMAN	P18583 homo sapien
39	77.5	6.8	252	1	VPHE_NPVAC	P24728 autographa
40	77.5	6.8	493	1	VE2_HPV19	P36786 human papil
41	77.5	6.8	673	1	LOX5_HUMAN	P09917 homo sapien
42	77.5	6.8	2404	1	SON_MOUSE	Q9qx47 mus musculu
43	77	6.7	77	1	PRT2_SEPOF	P80002 sepia offic
44	76	6.6	1286	1	CTRO_HUMAN	O14578 homo sapien
45	75.5	6.6	802	1	BCB2_ACEXY	O82860 acetobacter

ALIGNMENTS

RESULT 1
CORA_HPBV9
ID CORA_HPBV9 STANDARD; PRT; 214 AA.
AC P17099;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Core antigen.
GN C.
OS Hepatitis B virus (subtype adw / strain 991).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10410;
RN [1]
RP SEQUENCE FROM N.A.
RA Koechel H.G., Schueler A., Lottmann S., Thomssen R.;
RL Submitted (FEB-1990) to the EMBL/GenBank/DDBJ databases.
CC ---

CC ---
DR EMBL; X51970; CAA36232.1; -.
DR PIR; S10381; NKVLKS.
DR InterPro; IPR002006; Hepatitis_core.
DR Pfam; PF00906; Hepatitis_core; 1.
KW Core protein; Repeat.
FT REPEAT 193 200
FT REPEAT 201 208
SQ SEQUENCE 214 AA; 24722 MW; 2D668333EC5AFB8C CRC64;

Query Match 86.5%; Score 992.5; DB 1; Length 214;
Best Local Similarity 93.2%; Pred. No. 8.3e-82;
Matches 191; Conservative 2; Mismatches 5; Indels 7; Gaps 2;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||| |||||     |||||||:||||||||| |||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKIRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRR 190
           ||| || ||||||||:|||||||||||||||||||||||||||||||  |||||||
Db     130 LWFRISYLTFGRETVLEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRR 189

Qy     191 RTPSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 RTPSPRRRRSQSPRRRRSQSRESQC 214
```

RESULT 2

```
CORA_HPBVA
ID   CORA_HPBVA     STANDARD;      PRT;   211 AA.
AC   P24023;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Core antigen.
GN   C.
OS   Hepatitis B virus (strain alpha1).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10411;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90266476; PubMed=2345966;
RA   Tong S., Li J., Vitvitski L., Trepo C.;
RT   "Active hepatitis B virus replication in the presence of anti-HBe is
RT   associated with viral variants containing an inactive pre-C region.";
RL   Virology 176:596-603(1990).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M32138; -; NOT_ANNOTATED_CDS.
DR   PIR; A34773; NKVLA1.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT      183    190
FT   REPEAT      198    206
SQ   SEQUENCE   211 AA;  24208 MW;  B774AC72E65C75AB CRC64;

  Query Match             86.1%;  Score 988;  DB 1;  Length 211;
  Best Local Similarity   93.1%;  Pred. No. 2.1e-81;
  Matches  189;  Conservative    4;  Mismatches    4;  Indels    6;  Gaps    2;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWL-DMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 73

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||:|:||:|||| |||||      ||||||||||||||||||||
Db      74 EHCSPHHTALRQAILCWGDLITLSTWVGGNLEDP-----TSRDLVVSYVNTNMGLKFRQL 128

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 188

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||:|||| ||
Db     189 PSPRRRRSQSPRRRRTQSRESQC 211


RESULT 3
CORA_HPBVT
ID   CORA_HPBVT     STANDARD;      PRT;   212 AA.
AC   Q05495;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Core antigen.
GN   C.
OS   Hepatitis B virus (subtype adw4 / strain Brazil / isolate w4B).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=45410;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93346970; PubMed=8345355;
RA   Neumann H., Schaefer S., Yoshida C.F.T., Gaspar A.M.C., Repp R.,
RA   Gerlich W.H.;
RT   "Identification of a new hepatitis B virus (HBV) genotype from Brazil
RT   that expresses HBV surface antigen subtype adw4.";
RL   J. Gen. Virol. 74:1627-1632(1993).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X69798; CAA49452.1; -.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   DOMAIN      178    204       ARG-RICH.
FT   DOMAIN      184    203       3 X 5 AA REPEATS OF S-P-R-R-R.
FT   REPEAT      184    188       1.
FT   REPEAT      191    195       2.
FT   REPEAT      199    203       3.
SQ   SEQUENCE   212 AA;  24234 MW;  F832610DB7C36FD2 CRC64;

  Query Match             85.8%;  Score 985.5;  DB 1;  Length 212;
  Best Local Similarity   91.1%;  Pred. No. 3.5e-81;
  Matches  185;  Conservative    8;  Mismatches    5;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||:||||||||||||||||||||||||||:|||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGASVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||:|:|||||||||||||||:||| |||||      |:|||||:|||||||||| |||
Db      75 EHCTPNHTALRQAILCWGELMTLASWVGNNLEDP-----AARDLVVNYVNTNMGLKIRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVLEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||    ||
Db     190 PSPRRRRSQSPRRRRSQSPASQC 212


RESULT 4
CORA_HPBVY
ID   CORA_HPBVY     STANDARD;      PRT;   183 AA.
AC   P03146;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-APR-1990 (Rel. 14, Last annotation update)
DE   Core antigen.
GN   C.
OS   Hepatitis B virus (subtype ayw).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10418;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81012091; PubMed=399327;
RA   Galibert F., Mandart E., Fitoussi F., Tiollais P., Charnay P.;
RT   "Nucleotide sequence of the hepatitis B virus genome (subtype ayw)
RT   cloned in E. coli.";
RL   Nature 281:646-650(1979).
RN   [2]
RP   SEQUENCE FROM N.A. (CLONE PHB320).
RX   MEDLINE=85204397; PubMed=3996597;
RA   Bichko V., Pushko P., Dreilina D., Pumpen P., Gren E.;
RT   "Subtype ayw variant of hepatitis B virus. DNA primary structure
RT   analysis.";
RL   FEBS Lett. 185:208-212(1985).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------------
DR   EMBL; V01460; CAA24706.1; ALT_INIT.
DR   EMBL; X02496; -; NOT_ANNOTATED_CDS.
DR   PIR; A03711; NKVLAH.
DR   PIR; A03712; NKVLBH.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT       162    169
FT   REPEAT       170    177
FT   CONFLICT      33     33       T -> N (IN REF. 2).
FT   CONFLICT      80     80       A -> I (IN REF. 2).
SQ   SEQUENCE   183 AA;  21116 MW;  E0D9D9763F24E958 CRC64;

  Query Match             82.9%;  Score 951.5;  DB 1;  Length 183;
  Best Local Similarity   96.8%;  Pred. No. 3.2e-78;
  Matches  182;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           ||||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Db      61 CWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           ||||| ||
Db     176 SQSRESQC 183


RESULT 5
CORA_HPBVZ
ID   CORA_HPBVZ     STANDARD;      PRT;   183 AA.
AC   P03147;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-OCT-1989 (Rel. 12, Last annotation update)
DE   Core antigen.
GN   C.
OS   Hepatitis B virus (subtype adyw).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10419;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81012115; PubMed=399329;
RA   Pasek M., Goto T., Gilbert W., Zink B., Schaller H., McKay P.,
RA   Leadbetter G., Murray K.;
RT   "Hepatitis B virus genes and their expression in E. coli.";
RL   Nature 282:575-579(1979).
CC   ----------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------------
DR   EMBL; J02202; AAA45486.1; -.
DR   EMBL; A08967; CAA00816.1; -.
DR   PIR; B93217; NKVLA2.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT       162    169
FT   REPEAT       170    177
SQ   SEQUENCE   183 AA;  21042 MW;  545ED0E55527F26C CRC64;

  Query Match             81.2%;  Score 932.5;  DB 1;  Length 183;
  Best Local Similarity   93.6%;  Pred. No. 1.6e-76;
  Matches  176;  Conservative    5;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||:||||:||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAAALYRDALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||:|||||||| |||||      ||||||||||||:||||||||||||||||||||||
Db      61 CWGDLMTLATWVGTNLEDP-----ASRDLVVSYVNTNVGLKFRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           ||||| ||
Db     176 SQSRESQC 183


RESULT 6
CORA_HPBVJ
ID   CORA_HPBVJ     STANDARD;      PRT;   183 AA.
AC   P17391;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Core antigen.
OS   Hepatitis B virus (subtype adw / strain Japan/pJDW233).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10413;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89010694; PubMed=3171552;
RA   Okamoto H., Tsuda F., Sakugawa H., Sastrosoewignjo R.I., Imai M.,
RA   Miyakawa Y., Mayumi M.;
RT   "Typing hepatitis B virus by homology in nucleotide sequence:
RT   comparison of surface antigen subtypes.";
RL   J. Gen. Virol. 69:2575-2583(1988).
CC   ----------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------------
DR   EMBL; D00329; -; NOT_ANNOTATED_CDS.
DR   PIR; A28925; NKVLJ1.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT       162    169
FT   REPEAT       170    177
SQ   SEQUENCE   183 AA;  21224 MW;  9FDD6B5F5AF5E160 CRC64;

  Query Match             80.1%;  Score 919.5;  DB 1;  Length 183;
  Best Local Similarity   93.1%;  Pred. No. 2.3e-75;
  Matches  175;  Conservative    3;  Mismatches    5;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           |||||||||||||||||||||||||||||| ||||||||:||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTVSALYREALKSPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
```

```
          |||||||||||||| |||||     |||||||:||||||||| ||| ||||||||||||||||
Db      61 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKIRQLWWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           ||||| ||
Db     176 SQSRESQC 183


RESULT 7
CORA_HPBVO
ID   CORA_HPBVO     STANDARD;      PRT;   183 AA.
AC   P17392;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Core antigen.
OS   Hepatitis B virus (subtype adw / strain Okinawa/pODW282).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10415;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89010694; PubMed=3171552;
RA   Okamoto H., Tsuda F., Sakugawa H., Sastrosoewignjo R.I., Imai M.,
RA   Miyakawa Y., Mayumi M.;
RT   "Typing hepatitis B virus by homology in nucleotide sequence:
RT   comparison of surface antigen subtypes.";
RL   J. Gen. Virol. 69:2575-2583(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00330; -; NOT_ANNOTATED_CDS.
DR   PIR; B28925; NKVLJ2.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT      162    169
FT   REPEAT      170    177
SQ   SEQUENCE   183 AA;  21081 MW;  66BDB26331229300 CRC64;

  Query Match             80.1%;  Score 919.5;  DB 1;  Length 183;
  Best Local Similarity   93.1%;  Pred. No. 2.3e-75;
  Matches  175;  Conservative    3;  Mismatches    5;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGASVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||| |||||| |||||     |||:||||||| ||||| ||||||||||||||||||
Db      61 CWGELMNLATWVGSNLEDP-----ASRELVVSYVNVNMGLKIRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           ||||| ||
Db     176 SQSRESQC 183


RESULT 8
CORA_HPBVW
ID   CORA_HPBVW     STANDARD;      PRT;   185 AA.
AC   P03149;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-OCT-1989 (Rel. 12, Last annotation update)
DE   Core antigen.
GN   C.
OS   Hepatitis B virus (subtype adw).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=106821;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83168919; PubMed=6300776;
RA   Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
RT   "The complete nucleotide sequences of the cloned hepatitis B virus
RT   DNA; subtype adr and adw.";
RL   Nucleic Acids Res. 11:1747-1757(1983).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00866; -; NOT_ANNOTATED_CDS.
DR   PIR; C93460; NKVLA6.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT      164    171
FT   REPEAT      172    179
SQ   SEQUENCE   185 AA;  21394 MW;  B86A90D541BA70F9 CRC64;

  Query Match             80.1%;  Score 919.5;  DB 1;  Length 185;
  Best Local Similarity   93.2%;  Pred. No. 2.4e-75;
  Matches  177;  Conservative    3;  Mismatches    3;  Indels    7;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| ||:||     |||||||:||||||| ||||||||||||||||||
Db      61 CWGELMTLATWVGNNLQDP-----ASRDLVVNYVNTNMGLKIRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
           :||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRRSQSPRR 175

Qy     206 RRSQSREPQC 215
           ||||||| ||
Db     176 RRSQSRESQC 185


RESULT 9
CORA_HPBV4
ID   CORA_HPBV4     STANDARD;      PRT;   183 AA.
AC   P03150; P03151;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Core antigen.
GN   C.
OS   Hepatitis B virus (subtype adr4),
OS   Hepatitis B virus (subtype adr), and
OS   Hepatitis B virus (subtype adw / strain Indonesia/pIDW420).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10409, 106820, 10412;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ADR;
RX   MEDLINE=83168919; PubMed=6300776;
```

```
RA   Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
RT   "The complete nucleotide sequences of the cloned hepatitis B virus
RT   DNA; subtype adr and adw.";
RL   Nucleic Acids Res. 11:1747-1757(1983).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ADR4;
RX   MEDLINE=83246570; PubMed=6306594;
RA   Fujiyama A., Miyanohara A., Nozaki C., Yoneyama T., Ohtomo N.,
RA   Matsubara K.;
RT   "Cloning and structural analyses of hepatitis B virus DNAs, subtype
RT   adr.";
RL   Nucleic Acids Res. 11:4601-4610(1983).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ADW;
RX   MEDLINE=89010694; PubMed=3171552;
RA   Okamoto H., Tsuda F., Sakugawa H., Sastrosoewignjo R.I., Imai M.,
RA   Miyakawa Y., Mayumi M.;
RT   "Typing hepatitis B virus by homology in nucleotide sequence:
RT   comparison of surface antigen subtypes.";
RL   J. Gen. Virol. 69:2575-2583(1988).
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; V00867; -; NOT_ANNOTATED_CDS.
DR   EMBL; X01587; CAA25745.1; -.
DR   EMBL; D00331; -; NOT_ANNOTATED_CDS.
DR   PIR; A93480; NKVLA5.
DR   PIR; B93460; NKVLA4.
DR   PIR; C28925; NKVLJ3.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT      162    169
FT   REPEAT      170    177
SQ   SEQUENCE   183 AA;  21095 MW;  ED2DA1DB07FB596D CRC64;

  Query Match             80.0%;  Score 918.5;  DB 1;  Length 183;
  Best Local Similarity   92.6%;  Pred. No. 2.9e-75;
  Matches  174;  Conservative    4;  Mismatches    5;  Indels    5;  Gaps    1;

Qy         28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
              ||||||||||:|||||||||||||||:|||||||||||||||||||||||||||||||||
Db          1 MDIDPYKEFGASVELLSFLPSDFFPSIRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy         88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
              |||||| |||||| |||||     |||:||||||| ||||| ||||||||||||||||||
Db         61 CWGELMNLATWVGSNLEDP-----ASRELVVSYVNVNMGLKIRQLLWFHISCLTFGRETV 115

Qy        148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy        208 SQSREPQC 215
              ||||| ||
Db        176 SQSRESQC 183


RESULT 10
CORA_HPBVF
ID   CORA_HPBVF     STANDARD;      PRT;   195 AA.
AC   P29178;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   Core antigen.
GN   C.
OS   Hepatitis B virus (subtype adw2 variant sf).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=31515;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90169850; PubMed=2307406;
RA   Bhat R.A., Ulrich P.P., Vyas G.N.;
RT   "Molecular characterization of a new variant of hepatitis B virus in
RT   a persistently infected homosexual man.";
RL   Hepatology 11:271-276(1990).
DR   PIR; A37182; NKVLH3.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT      174    181
FT   REPEAT      182    189
SQ   SEQUENCE   195 AA;  22461 MW;  E2B166F879CB7CB7 CRC64;

  Query Match             79.5%;  Score 912.5;  DB 1;  Length 195;
  Best Local Similarity   91.1%;  Pred. No. 1.1e-74;
  Matches  174;  Conservative    6;  Mismatches    6;  Indels    5;  Gaps    1;

Qy         25 LWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQ 84
              |:|:||||||||||||||||||||||||||||||||||||||||:||| :||||||||||
Db         10 LFGLDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRESLESSDHCSPHHTALRQ 69

Qy         85 AILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGR 144
              |||| ||||||||||| |||||     |||||||:|||||||||| ||||||||||||||
Db         70 AILCWVELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKIRQLLWFHISCLTFGR 124

Qy        145 ETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPR 204
              |||:|||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        125 ETVLEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSPSPR 184

Qy        205 RRRSQSREPQC 215
              ||||||| ||
Db        185 RRRSQSRESQC 195


RESULT 11
CORA_HPBV2
ID   CORA_HPBV2     STANDARD;      PRT;   185 AA.
AC   P03148;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Core antigen.
GN   C.
OS   Hepatitis B virus (subtype adw2).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10408;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Valenzuela P., Quiroga M., Zaldivar J., Gray P., Rutter W.J.;
RL   (In) Field B.N., Jaenisch R., Fox C.F. (eds.);
RL   Animal virus genetics, pp.57-70, Academic Press, New York (1980).
DR   PIR; A94409; NKVLA3.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT      164    171
FT   REPEAT      172    179
SQ   SEQUENCE   185 AA;  21304 MW;  31F4DC338B507E19 CRC64;

  Query Match             79.4%;  Score 911.5;  DB 1;  Length 185;
  Best Local Similarity   92.6%;  Pred. No. 1.2e-74;
  Matches  176;  Conservative    3;  Mismatches    4;  Indels    7;  Gaps    2;

Qy         28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60
```

```
Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||     |||||||:|||||:||| ||||||||||||||||||
Db      61 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNVGLKIRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
           :|||||||||||||||||||||||||||||||||||||  |||||||||||||||| ||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRPSQSPRR 175

Qy     206 RRSQSREPQC 215
           ||||||| ||
Db     176 RRSQSRESQC 185


RESULT 12
CORA_HPBVL
ID   CORA_HPBVL     STANDARD;      PRT;   183 AA.
AC   P12901;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-APR-1990 (Rel. 14, Last annotation update)
DE   Core antigen.
GN   C.
OS   Hepatitis B virus (strain lsh / chimpanzee isolate).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10414;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88258473; PubMed=2838576;
RA   Vaudin M., Wolstenholme A.J., Tsiquaye K.N., Zuckerman A.J.,
RA   Harrison T.J.;
RT   "The complete nucleotide sequence of the genome of a hepatitis B
RT   virus isolated from a naturally infected chimpanzee.";
RL   J. Gen. Virol. 69:1383-1389(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00220; BAA00157.1; -.
DR   PIR; A28885; NKVLCP.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT      162    169
FT   REPEAT      170    177
SQ   SEQUENCE   183 AA;  20999 MW;  923DCB94A33FC0E8 CRC64;

  Query Match             78.4%;  Score 899.5;  DB 1;  Length 183;
  Best Local Similarity   91.5%;  Pred. No. 1.4e-73;
  Matches  172;  Conservative    5;  Mismatches    6;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           |||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPNHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||:||| |||||     |||: ||:|||||||| ||||||||||||||||||
Db      61 CWGELMTLASWVGNNLEDP-----ASREQVVNYVNTNMGLKIRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           |||    ||
Db     176 SQSPASQC 183


RESULT 13
```

```
CORA_HPBGS
ID   CORA_HPBGS     STANDARD;      PRT;   217 AA.
AC   P03153;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-JAN-1990 (Rel. 13, Last annotation update)
DE   Core antigen.
GN   C.
OS   Ground squirrel hepatitis virus (GSHV).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10406;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84267998; PubMed=6086950;
RA   Seeger C., Ganem D., Varmus H.E.;
RT   "Nucleotide sequence of an infectious molecularly cloned genome of
RT   ground squirrel hepatitis virus.";
RL   J. Virol. 51:367-375(1984).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02715; AAA46755.1; -.
DR   PIR; A03715; NKVLS.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT      196    203
FT   REPEAT      204    211
SQ   SEQUENCE   217 AA;  25189 MW;  DF489467355EC11A CRC64;

  Query Match             66.0%;  Score 757.5;  DB 1;  Length 217;
  Best Local Similarity   65.6%;  Pred. No. 8.7e-61;
  Matches  143;  Conservative   22;  Mismatches   42;  Indels   11;  Gaps    2;

Qy       3 LSSIFSRIGDPTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAS 62
           |  :|: :  |||||||||||||| ||||||||||:: :||:||| |||| :  |:|||:
Db       6 LCLVFACVPCPTVQASKLCLGWLWDMDIDPYKEFGSSYQLLNFLPLDFFPDLNALVDTAA 65

Qy      63 ALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVN 122
           ||| | |   |||||||||:|||::|| ||  | ||:  |         :  | ::| :||
Db      66 ALYEEELTGREHCSPHHTAIRQALVCWEELTRLITWMSENTT------EEVRRIIVDHVN 119

Qy     123 TNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVR 182
           ||| || |||||||||||: || |:|||||||||||  |||||||||||||| ||:|
Db     120 NTWGLKVRQTLWFHLSCLTFGQHTVQEFLVSFGVWIRTPAPYRPPNAPILSTLPEHTVIR 179

Qy     183 RRG-----RSPRRRTPSPRRRRSQSPRRRRSQSREPQC 215
           |||     |||||||||||||||||||||||||    |
Db     180 RRGGSRAARSPRRRTPSPRRRRSQSPRRRRSQSPASNC 217


RESULT 14
CORA_WHV1
ID   CORA_WHV1      STANDARD;      PRT;   188 AA.
AC   P03152;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Core antigen.
GN   C.
OS   Woodchuck hepatitis virus 1 (WHV 1),
OS   Woodchuck hepatitis virus 7 (WHV 7),
OS   Woodchuck hepatitis virus 59 (WHV 59), and
OS   Woodchuck hepatitis virus 8 (infectious clone) (WHV 8).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10430, 10432, 10431, 10434;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   SPECIES=Woodchuck hepatitis virus 1;
RX   MEDLINE=82216969; PubMed=7086958;
RA   Galibert F., Chen T.N., Mandart E.;
RT   "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:
RT   comparison with the hepatitis B virus sequence.";
RL   J. Virol. 41:51-65(1982).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Woodchuck hepatitis virus 7, and Woodchuck hepatitis virus 59;
RX   MEDLINE=88101359; PubMed=3336938;
RA   Cohen J.I., Miller R.H., Rosenblum B., Denniston K., Gerin J.L.,
RA   Purcell R.H.;
RT   "Sequence comparison of woodchuck hepatitis virus replicative forms
RT   shows conservation of the genome.";
RL   Virology 162:12-20(1988).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Woodchuck hepatitis virus 8 (infectious clone);
RX   MEDLINE=89184524; PubMed=2928306;
RA   Girones R., Cote P.J., Hornbuckle W.E., Tennant B.C., Gerin J.L.,
RA   Purcell R.H., Miller R.H.;
RT   "Complete nucleotide sequence of a molecular clone of woodchuck
RT   hepatitis virus that is infectious in the natural host.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:1846-1849(1989).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02442; AAA46761.1; -.
DR   EMBL; M18752; AAA46769.1; -.
DR   EMBL; M19183; AAA46765.1; -.
DR   EMBL; J04514; AAA46772.1; -.
DR   PIR; A03713; NKVLC.
DR   PIR; C32397; NKVLCT.
DR   PIR; C32397; NKVLC4.
DR   PIR; C32397; NKVLC3.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT      167    174
FT   REPEAT      175    182
SQ   SEQUENCE   188 AA;  21693 MW;  1F4454D0A7B7CE42 CRC64;

  Query Match             58.4%;  Score 671;  DB 1;  Length 188;
  Best Local Similarity   65.3%;  Pred. No. 4e-53;
  Matches  126;  Conservative   21;  Mismatches   36;  Indels   10;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||:: :||:||| |||| :  |:|||:||| | |   ||||||||||:|||::
Db       1 MDIDPYKEFGSSYQLLNFLPLDFFPDLNALVDTATALYEEELTGREHCSPHHTAIRQALV 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           || ||  |  |:  |:   :      | ::|:||   ||| || ||||:||||||: ||
Db      61 CWDELTKLIAWMSSNITSEQV-----RTIIVNHVNDTWGLKVRQSLWFHLSCLTFGQHTV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRG-----RSPRRRTPSPRRRRSQS 202
           |:||||||||||  |||||||||||||| ||:||||     ||||||||||||||||
Db     116 QEFLVSFGVWIRTPAPYRPPNAPILSTLPEHTVIRRRGGARASRSPRRRTPSPRRRRSQS 175

Qy     203 PRRRRSQSREPQC 215
           ||||||||    |
Db     176 PRRRRSQSPSANC 188


RESULT 15
CORA_WHV8
ID   CORA_WHV8      STANDARD;      PRT;   187 AA.
AC   P06433;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   01-OCT-1989 (Rel. 12, Last annotation update)
DE   Core antigen.
GN   C.
OS   Woodchuck hepatitis virus 8 (WHV 8).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10433;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86062931; PubMed=3855246;
RA   Kodama K., Ogasawara N., Yoshikawa H., Murakami S.;
RT   "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:
RT   evolutional relationship between hepadnaviruses.";
RL   J. Virol. 56:978-986(1985).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11082; AAA19185.1; -.
DR   PIR; A03714; NKVLC2.
DR   InterPro; IPR002006; Hepatitis_core.
DR   Pfam; PF00906; Hepatitis_core; 1.
KW   Core protein; Repeat.
FT   REPEAT      166    173
FT   REPEAT      174    181
SQ   SEQUENCE   187 AA;  21579 MW;  D4BC446FF7163165 CRC64;

  Query Match             57.1%;  Score 655.5;  DB 1;  Length 187;
  Best Local Similarity   64.6%;  Pred. No. 9.8e-52;
  Matches  124;  Conservative   21;  Mismatches   38;  Indels    9;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||:: :||:||| |||| :  |:|||:||| | |   ||||||||||:|||::
Db       1 MDIDPYKEFGSSYQLLNFLPLDFFPDLNALVDTATALYEEELTGREHCSPHHTAIRQALV 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           || ||  |  |:  |:   :      | ::|:||   ||| || ||||:||||||: ||
Db      61 CWDELTKLIAWMSSNITSEQV-----RTIIVNHVNDTWGLKVRQSLWFHLSCLTFGQHTV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRG----RSPRRRTPSPRRRRSQSP 203
           |:|||| ||||||  ||||||||||||||| ||:|| |    ||||||||||||||||||
Db     116 QEFLVSFVVWIRTPAPYRPPNAPILSTLPEHTVIRRGGARASRSPRRRTPSPRRRRSQSP 175

Qy     204 RRRRSQSREPQC 215
           |||||||    |
Db     176 RRRRSQSPSANC 187


Search completed: February  3, 2003, 09:24:20
Job time : 9.86585 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:57 ; Search time 14.2334 Seconds
(without alignments)
1452.136 Million cell updates/sec

Title: US-09-890-752A-2
Perfect score: 1148
Sequence: 1 MPLSSIFSRIGDPTVQASKL.....RRRRSQSPRRRRSQSREPQC 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1040.5	90.6	212	1	NKVLAH	e antigen precurso
2	1038.5	90.5	212	2	S53204	e antigen precurso
3	1037.5	90.4	212	2	S53211	e antigen precurso
4	1037.5	90.4	212	2	S32204	e antigen precurso
5	1037.5	90.4	212	2	S20750	e antigen precurso
6	1035.5	90.2	212	2	S53216	e antigen precurso
7	1035.5	90.2	212	2	S53272	e antigen precurso
8	1034.5	90.1	212	2	S53225	e antigen precurso
9	1034.5	90.1	212	2	S53274	e antigen precurso
10	1034.5	90.1	212	2	S53163	e antigen precurso
11	1031.5	89.9	212	2	S53251	e antigen precurso
12	1031.5	89.9	212	2	S20746	e antigen precurso
13	1030.5	89.8	212	2	S53159	e antigen precurso
14	1030.5	89.8	212	2	S53200	e antigen precurso
15	1030.5	89.8	212	2	S53202	e antigen precurso
16	1029.5	89.7	212	2	S53281	e antigen precurso
17	1029.5	89.7	212	2	S53198	e antigen precurso
18	1028.5	89.6	212	2	S53223	e antigen precurso
19	1028.5	89.6	212	2	S53227	e antigen precurso
20	1025.5	89.3	212	2	S53253	e antigen precurso
21	1023.5	89.2	212	1	NKVLBH	e antigen precurso
22	1022.5	89.1	212	2	S53240	e antigen precurso
23	1020.5	88.9	212	2	S53242	e antigen precurso
24	1016.5	88.5	212	2	S53257	e antigen precurso
25	1015.5	88.5	212	2	S53279	e antigen precurso
26	1014.5	88.4	212	2	S53229	e antigen precurso
27	1014.5	88.4	212	2	S25651	e antigen precurso
28	1012.5	88.2	212	2	S53255	e antigen precurso
29	1011.5	88.1	212	2	S53238	e antigen precurso
30	1011.5	88.1	214	2	S01405	core antigen - hep
31	1008.5	87.8	212	1	NKVLJ1	e antigen precurso
32	1008.5	87.8	212	1	NKVLJ2	e antigen precurso
33	1008.5	87.8	214	1	NKVLA6	e antigen precurso
34	1007.5	87.8	212	1	NKVLA4	e antigen precurso
35	1007.5	87.8	212	2	S53236	e antigen precurso
36	1007.5	87.8	212	2	S53157	e antigen precurso
37	1006.5	87.7	214	2	S47409	e antigen precurso
38	1002.5	87.3	214	2	S33686	e antigen precurso
39	1001.5	87.2	214	1	NKVLA3	e antigen precurso
40	999.5	87.1	212	2	S53161	e antigen precurso
41	999.5	87.1	359	2	S04570	gene X/C fusion pr
42	992.5	86.5	214	1	NKVLKS	e antigen precurso
43	988	86.1	211	1	NKVLA1	e antigen precurso
44	985.5	85.8	212	2	JQ2227	e antigen precurso
45	964.5	84.0	212	2	S67504	e antigen precurso

ALIGNMENTS

RESULT 1
NKVLAH
e antigen precursor / core antigen - hepatitis B virus (subtype ayw4, isolate hb321 and
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw4, isolate hb321; isolate patient Ferracuti'83; isolate patient Ca
cheri'83
C;Date: 18-Dec-1981 #sequence_revision 08-Nov-1996 #text_change 16-Jul-1999
C;Accession: S47405; S53191; S53209; S53234; S53264; S53249; S53262; S53277; A03711
R;Pluciennicza k, A.
submitted to the EMBL Data Library, August 1994
A;Description: Molecular cloning and sequencing of two complete genomes of polish isolat
A;Reference number: S47404
A;Accession: S47405
A;Molecule type: DNA
A;Residues: 1-212 <PLU>
A;Cross-references: EMBL:Z35716; NID:g527435; PIDN:CAA84786.1; PID:g527437
A;Experimental source: subtype ayw4, isolate hb321
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53191
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85283; NID:g736088; PIDN:CAA59593.1; PID:g736090
A;Experimental source: isolate patient Ferracuti'83
A;Accession: S53209
A;Molecule type: DNA
A;Residues: 1-212 <LA2>
A;Cross-references: EMBL:X85290; NID:g736114; PIDN:CAA59609.1; PID:g736116
A;Experimental source: isolate patient Castag'83
A;Accession: S53234
A;Molecule type: DNA
A;Residues: 1-212 <LA3>
A;Cross-references: EMBL:X85300; NID:g736150; PIDN:CAA59631.1; PID:g736152
A;Experimental source: isolate patient Sanna'84
A;Accession: S53264
A;Molecule type: DNA
A;Residues: 1-212 <LA4>
A;Cross-references: EMBL:X85313; NID:g736194; PIDN:CAA59659.1; PID:g736196
A;Experimental source: isolate patient Licheri-1'85
A;Accession: S53249
A;Molecule type: DNA
A;Residues: 1-212 <LA5>
A;Cross-references: EMBL:X85306; NID:g736172; PIDN:CAA59644.1; PID:g736174
A;Experimental source: isolate patient Flore-1'86
A;Accession: S53262
A;Molecule type: DNA
A;Residues: 1-212 <LA6>
A;Cross-references: EMBL:X85312; NID:g736191; PIDN:CAA59657.1; PID:g736193
A;Experimental source: isolate patient Licheri'83

A;Accession: S53277
A;Molecule type: DNA
A;Residues: 30-212 <LA7>
A;Cross-references: EMBL:X85317; NID:g736211; PIDN:CAA59669.1; PID:g736214
A;Experimental source: patient Giordo-2'86
A;Note: due to a stop codon between the alternative initiators the e antigen precursor c
R;Galibert, F.; Mandart, E.; Fitoussi, F.; Tiollais, P.; Charnay, P.
Nature 281, 646-650, 1979
A;Title: Nucleotide sequence of the hepatitis B virus genome (subtype ayw) in E. coli.
A;Reference number: A93214; MUID:81012091; PMID:399327
A;Accession: A03711
A;Molecule type: DNA
A;Residues: 1-212 <GAL>
A;Cross-references: GB:J02203; NID:g329640; PIDN:AAA45489.1; PID:g329642
A;Experimental source: subtype ayw
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match 90.6%; Score 1040.5; DB 1; Length 212;
Best Local Similarity 97.0%; Pred. No. 2e-82;
Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 2
S53204
e antigen precursor / core antigen - hepatitis B virus
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53204; S53283
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53204
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85288; NID:g736107; PIDN:CAA59604.1; PID:g736109
A;Experimental source: isolate patient Castaa-1'85
A;Accession: S53283
A;Molecule type: DNA
A;Residues: 1-14,'A',16-57,'D',59-61,'N',63-128,'R',130-209,'S',211-212 <LA2>
A;Cross-references: EMBL:X85320; NID:g736221; PIDN:CAA59675.1; PID:g736223
A;Experimental source: isolate patient Galistru'85
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match 90.5%; Score 1038.5; DB 2; Length 212;
Best Local Similarity 96.6%; Pred. No. 3e-82;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRALLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||:
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQI 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSREPQC 212
```

RESULT 3
S53211
e antigen precursor / core antigen - hepatitis B virus (isolate patient Castag-1'85 and
N;Alternate names: HBc antigen; HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Castag-1'85; isolate patient Ferracuti-1'89
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53211; S53197
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53211
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85291; NID:g736117; PIDN:CAA59611.1; PID:g736119
A;Experimental source: isolate patient Castag-1'85
A;Accession: S53197
A;Molecule type: DNA
A;Residues: 30-212 <LAW>
A;Cross-references: EMBL:X85284; NID:g736095; PIDN:CAA59596.1; PID:g736098
A;Experimental source: isolate patient Ferracuti-1'89
A;Note: due to a stop codon between the alternative initiatiors the e antigen precursor
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match 90.4%; Score 1037.5; DB 2; Length 212;
Best Local Similarity 96.6%; Pred. No. 3.6e-82;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||:||||||||||||     ||||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGDLMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||| ||
```

```
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212

RESULT 4
S32204
e antigen precursor / core antigen - hepatitis B virus (subtype ayw, patients C1000 and
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Oct-2000
C;Accession: S32204; S53207
R;Preisler-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Rasenack, J.
submitted to the EMBL Data Library, March 1993
A;Description: Identification and sequence analysis of hepatitis B virus DNA in immunolo
A;Reference number: S32202
A;Accession: S32204
A;Molecule type: DNA
A;Residues: 1-212 <PRE>
A;Cross-references: EMBL:X72702; NID:g288927; PIDN:CAA51257.1; PID:g288930
A;Experimental source: subtype ayw, patient C1000
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53207
A;Molecule type: DNA
A;Residues: 30-212 <LAI>
A;Cross-references: EMBL:X85289; NID:g736110; PIDN:CAA59607.1; PID:g736113
A;Experimental source: isolate patient Castaa-2'87
A;Note: due to a stop codon mutation between the alternative initiatiors the e antigen p
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

  Query Match             90.4%;  Score 1037.5;  DB 2;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 3.6e-82;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy     13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db     15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy     73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          ||||||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db     75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy    133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          |||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212

RESULT 5
S20750
e antigen precursor / core antigen - hepatitis B virus (subtype ayw, patient CI)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient CI
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S20750
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submitted to the EMBL Data Library, March 1992
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati
A;Reference number: S20745
A;Accession: S20750
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X65258; NID:g59434; PIDN:CAA46354.1; PID:g59436
A;Experimental source: subtype ayw, patient CI
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

  Query Match             90.4%;  Score 1037.5;  DB 2;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 3.6e-82;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy     13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy     73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          ||||||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db     75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy    133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
          ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db    130 LWFHISCLTFGRETVIEYLVAFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          |||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212

RESULT 6
S53216
e antigen precursor / core antigen - hepatitis B virus (isolate patient Castag'3)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Castag'3
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53216
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53216
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85293; NID:g736124; PIDN:CAA59616.1; PID:g736126
A;Experimental source: isolate patient Castag'3
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

  Query Match             90.2%;  Score 1035.5;  DB 2;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 5.4e-82;
  Matches  196;  Conservative    0;  Mismatches    2;  Indels    5;  Gaps    1;

Qy     13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDNASALYREALESP 74

Qy     73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          ||||||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db     75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy    133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          |||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 7
S53272
e antigen precursor / core antigen - hepatitis B virus (isolate patient Licheri-3'90)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Licheri-3'90
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53272
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53272
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85315; NID:g736205; PIDN:CAA59664.1; PID:g736207
A;Experimental source: isolate patient Licheri-3'90
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

  Query Match           90.2%;  Score 1035.5;  DB 2;  Length 212;
  Best Local Similarity  96.6%;  Pred. No. 5.4e-82;
  Matches  196;  Conservative    0;  Mismatches    2;  Indels    5;  Gaps    1;

Qy     13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy     73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db     75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy    133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 LWFHICCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          |||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 8
S53225
e antigen precursor / core antigen - hepatitis B virus (isolate patient Chighine-2'86)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Chighine-2'86
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53225
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53225
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85296; NID:g736137; PIDN:CAA59622.1; PID:g736139
A;Experimental source: isolate patient Chighine-2'86
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

  Query Match           90.1%;  Score 1034.5;  DB 2;  Length 212;
  Best Local Similarity  96.1%;  Pred. No. 6.6e-82;
  Matches  195;  Conservative    2;  Mismatches    1;  Indels    5;  Gaps    1;

Qy     13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy     73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          ||||||||||||||||:|||||||||||||     |||||||||||||||||||:||
Db     75 EHCSPHHTALRQAILCWGDLMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFKQL 129

Qy    133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          |||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 9
S53274
e antigen precursor / core antigen - hepatitis B virus (isolate patient Giordo'84)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Giordo'84
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53274
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53274
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85316; NID:g736208; PIDN:CAA59666.1; PID:g736210
A;Experimental source: isolate patient patient Giordo'84
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

  Query Match           90.1%;  Score 1034.5;  DB 2;  Length 212;
  Best Local Similarity  96.6%;  Pred. No. 6.6e-82;
  Matches  196;  Conservative    0;  Mismatches    2;  Indels    5;  Gaps    1;

Qy     13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy     73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db     75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy    133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILLTLPETTVVRRRGRSPRRRT 189

Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          |||||||||||||||||||| ||
```

```
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 10
S53163
e antigen precursor / core antigen - hepatitis B virus (isolate patient Vittorina'92)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Vittorina'92
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53163
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53163
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85256; NID:g736050; PIDN:CAA59519.1; PID:g736052
A;Experimental source: isolate patient Vittorina'92
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

  Query Match             90.1%;  Score 1034.5;  DB 2;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 6.6e-82;
  Matches  196;  Conservative    0;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLSTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 11
S53251
e antigen precursor / core antigen - hepatitis B virus (isolate patient Flore'85)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Flore'85
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53251
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53251
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85307; NID:g736175; PIDN:CAA59646.1; PID:g736177
A;Experimental source: isolate patient Flore'85
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

  Query Match             89.9%;  Score 1031.5;  DB 2;  Length 212;
  Best Local Similarity   96.1%;  Pred. No. 1.2e-81;
  Matches  195;  Conservative    1;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||:| ||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGDLTTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 12
S20746
e antigen precursor / core antigen - hepatitis B virus (subtype ayw, patient C)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, patient C
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S20746
R;Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submitted to the EMBL Data Library, March 1992
A;Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negativ
A;Reference number: S20745
A;Accession: S20746
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X65257; NID:g59429; PIDN:CAA46350.1; PID:g59431
A;Experimental source: subtype ayw, patient C
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

  Query Match             89.9%;  Score 1031.5;  DB 2;  Length 212;
  Best Local Similarity   96.1%;  Pred. No. 1.2e-81;
  Matches  195;  Conservative    1;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLIFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 13
```

S53159
e antigen precursor / core antigen - hepatitis B virus (isolate patient Garrucciu'90)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Garrucciu'90
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53159
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53159
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85272; NID:g736044; PIDN:CAA59563.1; PID:g736046
A;Experimental source: isolate patient Garrucciu'90
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match 89.8%; Score 1030.5; DB 2; Length 212;
Best Local Similarity 95.6%; Pred. No. 1.5e-81;
Matches 194; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||:|||||||||||||||     |:|||||||||||:|||||||
Db      75 EHCSPHHTALRQAILCWGDLMTLATWVGVNLEDP-----ATRDLVVSYVNTNVGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 14
S53200
e antigen precursor / core antigen - hepatitis B virus (isolate patient Ferracuti-3'91)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Ferracuti-3'91
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53200
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53200
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85286; NID:g736101; PIDN:CAA59600.1; PID:g736103
A;Experimental source: isolate patient Ferracuti-3'91
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match 89.8%; Score 1030.5; DB 2; Length 212;
Best Local Similarity 96.1%; Pred. No. 1.5e-81;
Matches 195; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYKPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRGSQC 212
```

RESULT 15
S53202
e antigen precursor / core antigen - hepatitis B virus (isolate patient Castaa'83)
N;Alternate names: HBe antigen precursor / HBc antigen; pre-C/C antigen
N;Contains: core antigen; e antigen
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Castaa'83
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C;Accession: S53202
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53202
A;Molecule type: DNA
A;Residues: 1-212 <LAI>
A;Cross-references: EMBL:X85287; NID:g736104; PIDN:CAA59602.1; PID:g736106
A;Experimental source: isolate patient Castaa'83
C;Genetics:
A;Gene: C
C;Superfamily: hepatitis B virus core antigen
C;Keywords: alternative initiators; core protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-212/Product: core antigen #status predicted <CAG>
F;30-178/Product: e antigen #status predicted <EAG>
F;179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match 89.8%; Score 1030.5; DB 2; Length 212;
Best Local Similarity 96.1%; Pred. No. 1.5e-81;
Matches 195; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRALLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     |||||||||||||||||||:
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQI 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

Search completed: February 3, 2003, 09:26:28
Job time : 15.2334 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:24:27 ; Search time 17.6045 Seconds
(without alignments)
988.252 Million cell updates/sec

Title: US-09-890-752A-2
Perfect score: 1148
Sequence: 1 MPLSSIFSRIGDPTVQASKL.....RRRRSQSPRRRRSQSREPQC 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 404799 seqs, 80919614 residues

Total number of hits satisfying chosen parameters: 404799

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	951.5	82.9	183	6	US-10-274-616-1	Sequence 1, Appli
2	932.5	81.2	183	6	US-10-274-616-4	Sequence 4, Appli
3	919.5	80.1	185	6	US-10-274-616-2	Sequence 2, Appli
4	911.5	79.4	185	6	US-10-274-616-3	Sequence 3, Appli
5	901.5	78.5	183	6	US-10-226-007-31	Sequence 31, Appl
6	754.5	65.7	217	6	US-10-274-616-6	Sequence 6, Appli
7	722.5	62.9	169	1	PCT-US02-17536-17	Sequence 17, Appl
8	722.5	62.9	169	1	PCT-US02-17536-18	Sequence 18, Appl
9	661	57.6	183	6	US-10-274-616-5	Sequence 5, Appli
10	114	9.9	20	5	US-09-863-054-21	Sequence 21, Appl
11	113.5	9.9	312	6	US-10-218-140-1758	Sequence 1758, Ap
12	110	9.6	20	5	US-09-863-054-19	Sequence 19, Appl
13	110	9.6	20	6	US-10-226-007-714	Sequence 714, App
14	109	9.5	20	6	US-10-226-007-646	Sequence 646, App
15	109	9.5	20	6	US-10-226-007-679	Sequence 679, App
16	109	9.5	20	6	US-10-226-007-688	Sequence 688, App
17	109	9.5	20	6	US-10-226-007-696	Sequence 696, App
18	109	9.5	20	6	US-10-226-007-703	Sequence 703, App
19	109	9.5	20	6	US-10-226-007-709	Sequence 709, App
20	108	9.4	20	6	US-10-226-007-669	Sequence 669, App
21	108	9.4	20	6	US-10-226-007-718	Sequence 718, App
22	106	9.2	20	6	US-10-226-007-658	Sequence 658, App
23	105	9.1	19	6	US-10-226-007-678	Sequence 678, App
24	105	9.1	19	6	US-10-226-007-687	Sequence 687, App
25	105	9.1	19	6	US-10-226-007-695	Sequence 695, App
26	105	9.1	19	6	US-10-226-007-708	Sequence 708, App
27	105	9.1	20	6	US-10-226-007-721	Sequence 721, App
28	104	9.1	19	6	US-10-226-007-713	Sequence 713, App
29	104	9.1	20	5	US-09-863-054-17	Sequence 17, Appl
30	104	9.1	20	6	US-10-226-007-607	Sequence 607, App
31	104	9.1	20	6	US-10-226-007-620	Sequence 620, App
32	103.5	9.0	32	6	US-10-326-908-12	Sequence 12, Appl
33	103	9.0	19	6	US-10-226-007-668	Sequence 668, App
34	103	9.0	20	6	US-10-226-007-594	Sequence 594, App
35	103	9.0	20	6	US-10-226-007-633	Sequence 633, App
36	103	9.0	20	6	US-10-226-007-723	Sequence 723, App
37	102.5	8.9	24	6	US-10-326-908-11	Sequence 11, Appl
38	102	8.9	19	6	US-10-226-007-645	Sequence 645, App
39	102	8.9	19	6	US-10-226-007-657	Sequence 657, App
40	102	8.9	19	6	US-10-226-007-702	Sequence 702, App
41	101	8.8	18	6	US-10-226-007-677	Sequence 677, App
42	101	8.8	18	6	US-10-226-007-686	Sequence 686, App
43	101	8.8	19	6	US-10-226-007-717	Sequence 717, App
44	100	8.7	20	5	US-09-865-294A-20	Sequence 20, Appl
45	100	8.7	20	6	US-10-226-007-724	Sequence 724, App

ALIGNMENTS

RESULT 1
US-10-274-616-1
; Sequence 1, Application US/10274616
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
; FILE REFERENCE: ICC 127.0 4564/88545
; CURRENT APPLICATION NUMBER: US/10/274,616
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: 09/930,915
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 10/080,299
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-274-616-1

Query Match 82.9%; Score 951.5; DB 6; Length 183;
Best Local Similarity 96.8%; Pred. No. 7.4e-62;
Matches 182; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

```
Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||||||||     ||||||||||||||||||||||||||||||||||||
Db      61 CWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           ||||| ||
Db     176 SQSRESQC 183
```

RESULT 2
US-10-274-616-4
; Sequence 4, Application US/10274616
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
; FILE REFERENCE: ICC 127.0 4564/88545

```
;  CURRENT APPLICATION NUMBER: US/10/274,616
;  CURRENT FILING DATE:  2002-10-21
;  PRIOR APPLICATION NUMBER: 09/930,915
;  PRIOR FILING DATE: 2001-08-15
;  PRIOR APPLICATION NUMBER: 10/080,299
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 183
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-274-616-4

  Query Match             81.2%;  Score 932.5;  DB 6;  Length 183;
  Best Local Similarity   93.6%;  Pred. No. 1.8e-60;
  Matches  176;  Conservative    5;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||:||||:||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAAALYRDALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||:|||||||||| |||||     |||||||||||||:||||||||||||||||||||
Db      61 CWGDLMTLATWVGTNLEDP-----ASRDLVVSYVNTNVGLKFRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           ||||| ||
Db     176 SQSRESQC 183


RESULT 3
US-10-274-616-2
; Sequence 2, Application US/10274616
; GENERAL INFORMATION:
;  APPLICANT: Birkett, Ashley J.
;  TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
;  FILE REFERENCE: ICC 127.0 4564/88545
;  CURRENT APPLICATION NUMBER: US/10/274,616
;  CURRENT FILING DATE:  2002-10-21
;  PRIOR APPLICATION NUMBER: 09/930,915
;  PRIOR FILING DATE: 2001-08-15
;  PRIOR APPLICATION NUMBER: 10/080,299
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 185
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-274-616-2

  Query Match             80.1%;  Score 919.5;  DB 6;  Length 185;
  Best Local Similarity   93.2%;  Pred. No. 1.5e-59;
  Matches  177;  Conservative    3;  Mismatches    3;  Indels    7;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| ||:||     |||||||:||||||||| ||||||||||||||||||
Db      61 CWGELMTLATWVGNNLQDP-----ASRDLVVNYVNTNMGLKIRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
           :|||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRRSQSPRR 175

Qy     206 RRSQSREPQC 215
           ||||||| ||
Db     176 RRSQSRESQC 185


RESULT 4
US-10-274-616-3
; Sequence 3, Application US/10274616
; GENERAL INFORMATION:
;  APPLICANT: Birkett, Ashley J.
;  TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
;  FILE REFERENCE: ICC 127.0 4564/88545
;  CURRENT APPLICATION NUMBER: US/10/274,616
;  CURRENT FILING DATE:  2002-10-21
;  PRIOR APPLICATION NUMBER: 09/930,915
;  PRIOR FILING DATE: 2001-08-15
;  PRIOR APPLICATION NUMBER: 10/080,299
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 185
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-274-616-3

  Query Match             79.4%;  Score 911.5;  DB 6;  Length 185;
  Best Local Similarity   92.6%;  Pred. No. 5.8e-59;
  Matches  176;  Conservative    3;  Mismatches    4;  Indels    7;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||     |||||||:|||:||| ||||||||||||||||||
Db      61 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNVGLKIRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
           :|||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRRPSQSPRR 175

Qy     206 RRSQSREPQC 215
           ||||||| ||
Db     176 RRSQSRESQC 185


RESULT 5
US-10-226-007-31
; Sequence 31, Application US/10226007
; GENERAL INFORMATION:
;  APPLICANT: Myriad Genetics, Inc.
;  APPLICANT:  Morham, Scott
;  APPLICANT:  Zavitz, Kenton
;  APPLICANT:  Hobden, Adrian
;  TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
;  FILE REFERENCE: 5005.01
;  CURRENT APPLICATION NUMBER: US/10/226,007
;  CURRENT FILING DATE:  2002-11-15
;  PRIOR APPLICATION NUMBER: US 60/313,883
;  PRIOR FILING DATE: 2001-08-21
;  NUMBER OF SEQ ID NOS: 1673
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
;   LENGTH: 183
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-226-007-31

  Query Match             78.5%;  Score 901.5;  DB 6;  Length 183;
  Best Local Similarity   90.4%;  Pred. No. 3.1e-58;
  Matches  170;  Conservative    8;  Mismatches    5;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
```

```
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAAALFREALESPEHCTPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
Db      61 CWGELMSLATWVGVNLED-----QASRDLVVRYVNDHMGIKFRQLLWFHISCLTFGRDTV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
Db     116 IEYLVSFGVWIRTPPPYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
Db     176 TQSRESQC 183


RESULT 6
US-10-274-616-6
; Sequence 6, Application US/10274616
; GENERAL INFORMATION:
;  APPLICANT: Birkett, Ashley J.
;  TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
;  FILE REFERENCE: ICC 127.0 4564/88545
;  CURRENT APPLICATION NUMBER: US/10/274,616
;  CURRENT FILING DATE:  2002-10-21
;  PRIOR APPLICATION NUMBER: 09/930,915
;  PRIOR FILING DATE: 2001-08-15
;  PRIOR APPLICATION NUMBER: 10/080,299
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 217
;   TYPE: PRT
;   ORGANISM: Spermophilus variegatus
US-10-274-616-6

  Query Match           65.7%;  Score 754.5;  DB 6;  Length 217;
  Best Local Similarity  65.6%;  Pred. No. 1.5e-47;
  Matches  143;  Conservative   21;  Mismatches   43;  Indels   11;  Gaps    2;

Qy       3 LSSIFSRIGDPTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAS 62
Db       6 LCLVFACVPCPTVQASKLCLGWLWDMDIDPYKEFGSSYQLLNFLPLDFFPDLNALVDTAA 65

Qy      63 ALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVN 122
Db      66 ALYEEELTGREHCSPHHTAIRQALVCWEELTRLITWMSENTT------EEVRRIIVDHVN 119

Qy     123 TNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVR 182
Db     120 NTWGLKVRQTLWFHLSCLTFGGHTVQEFLVSFGVWIRTPAPYRPPNAPILSTLPEHTVIR 179

Qy     183 RRG-----RSPRRRTPSPRRRRSQSPRRRRSQSREPQC 215
Db     180 RRGGSRAARSPRRRTPSPRRRRSQSPRRRRSQSPASNC 217


RESULT 7
PCT-US02-17536-17
; Sequence 17, Application PC/TUS0217536
; GENERAL INFORMATION:
;  APPLICANT: Holoshitz, Joseph
;  APPLICANT:  Ling, Song
;  TITLE OF INVENTION: Methods and Compositions for the Treatment of Diseases Associated
;  TITLE OF INVENTION:  Signal Transduction Aberrations
;  FILE REFERENCE: UM-07135
;  CURRENT APPLICATION NUMBER: PCT/US02/17536
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: 60/295,691
;  PRIOR FILING DATE: 2001-06-04
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
;   LENGTH: 169
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
PCT-US02-17536-17

  Query Match           62.9%;  Score 722.5;  DB 1;  Length 169;
  Best Local Similarity  87.7%;  Pred. No. 2.6e-45;
  Matches  142;  Conservative    1;  Mismatches    6;  Indels   13;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDNASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDP----EFRGDA---------SRDLVVSYVNTNMGLKFRQLLW 134
Db      61 CWGELMTLATWVGGNLEDHKDLLEQKRAAVDTYCVDPISRDLVVSYVNTNMGLKFRQLLW 120

Qy     135 FHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLP 176
Db     121 FHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLP 162


RESULT 8
PCT-US02-17536-18
; Sequence 18, Application PC/TUS0217536
; GENERAL INFORMATION:
;  APPLICANT: Holoshitz, Joseph
;  APPLICANT:  Ling, Song
;  TITLE OF INVENTION: Methods and Compositions for the Treatment of Diseases Associated
;  TITLE OF INVENTION:  Signal Transduction Aberrations
;  FILE REFERENCE: UM-07135
;  CURRENT APPLICATION NUMBER: PCT/US02/17536
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: 60/295,691
;  PRIOR FILING DATE: 2001-06-04
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 169
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
PCT-US02-17536-18

  Query Match           62.9%;  Score 722.5;  DB 1;  Length 169;
  Best Local Similarity  87.0%;  Pred. No. 2.6e-45;
  Matches  141;  Conservative    1;  Mismatches    7;  Indels   13;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDNASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLED-PEFRGD-----------ASRDLVVSYVNTNMGLKFRQLLW 134
Db      61 CWGELMTLATWVGGNLEDHKDILEDERAAVDTYCVDPISRDLVVSYVNTNMGLKFRQLLW 120

Qy     135 FHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLP 176
Db     121 FHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLP 162


RESULT 9
US-10-274-616-5
; Sequence 5, Application US/10274616
; GENERAL INFORMATION:
;  APPLICANT: Birkett, Ashley J.
;  TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
;  FILE REFERENCE: ICC 127.0 4564/88545
;  CURRENT APPLICATION NUMBER: US/10/274,616
```

```
;  CURRENT FILING DATE:  2002-10-21
;  PRIOR APPLICATION NUMBER: 09/930,915
;  PRIOR FILING DATE: 2001-08-15
;  PRIOR APPLICATION NUMBER: 10/080,299
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 183
;   TYPE: PRT
;   ORGANISM: Marmota monax
US-10-274-616-5

  Query Match             57.6%;  Score 661;  DB 6;  Length 183;
  Best Local Similarity   66.3%;  Pred. No. 7.8e-41;
  Matches  124;  Conservative   21;  Mismatches   32;  Indels   10;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||:: :||:||| |||| :  |:|||:||| | |   |||||||||:|||::
Db       1 MDIDPYKEFGSSYQLLNFLPLDFFPDLNALVDTATALYEEELTGREHCSPHHTAIRQALV 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           || ||  |  |:  |:   :      | ::|::||   ||| || ||||:||||||: ||
Db      61 CWDELTKLIAWMSSNITSEQV-----RTIIVNHVNDTWGLKVRQSLWFHLSCLTFGQHTV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRG-----RSPRRRTPSPRRRRSQS 202
           |:|||||||||||  |||||||||||||| ||:||||     |||||||||||||||||
Db     116 QEFLVSFGVWIRTPAPYRPPNAPILSTLPEHTVIRRRGGARASRSPRRRTPSPRRRRSQS 175

Qy     203 PRRRRSQ 209
           |||||||
Db     176 PRRRRSQ 182


RESULT 10
US-09-863-054-21
; Sequence 21, Application US/09863054
;   GENERAL INFORMATION:
;        APPLICANT: Chisari, Francis V.
;        TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
;                            Lymphocyte Responses to Hepatitis B Virus
;        NUMBER OF SEQUENCES: 57
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/863,054
;             FILING DATE: 21-May-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 07/935,898
;             FILING DATE: 26-AUG-1992
;             APPLICATION NUMBER: US 08/024,120
;             FILING DATE: 26-FEB-1993
;             APPLICATION NUMBER: US 08/396,283
;             FILING DATE: 27-FEB-1995
;             APPLICATION NUMBER: US 08/463,486
;             FILING DATE: 05-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lockyer, Jean M.
;             REGISTRATION NUMBER: 44,879
;             REFERENCE/DOCKET NUMBER: 014740-000421US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 576-0200
;             TELEFAX: (415) 576-0300
;   INFORMATION FOR SEQ ID NO: 21:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 20 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-863-054-21

  Query Match             9.9%;  Score 114;  DB 5;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.044;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    152 VSFGVWIRTPPAYRPPNAPI 171
          ||||||||||||||||||||
Db      1 VSFGVWIRTPPAYRPPNAPI 20


RESULT 11
US-10-218-140-1758
; Sequence 1758, Application US/10218140
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin D.
;  APPLICANT:  Shimkets, Richard A.
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
;  TITLE OF INVENTION:  ENCODED THEREBY
;  FILE REFERENCE: 15966-543 CON
;  CURRENT APPLICATION NUMBER: US/10/218,140
;  CURRENT FILING DATE:  2002-08-12
;  PRIOR APPLICATION NUMBER: 09/540,763
;  PRIOR FILING DATE: 2000-03-30
;  PRIOR APPLICATION NUMBER: 60/127,728
;  PRIOR FILING DATE: 1999-04-05
;  PRIOR APPLICATION NUMBER: 60/127,636
;  PRIOR FILING DATE: 1999-04-02
;  PRIOR APPLICATION NUMBER: 60/127,607
;  PRIOR FILING DATE: 1999-03-31
;  NUMBER OF SEQ ID NOS: 6322
;  SOFTWARE: CuraNator Version 1.0
; SEQ ID NO 1758
;   LENGTH: 312
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-140-1758

  Query Match             9.9%;  Score 113.5;  DB 6;  Length 312;
  Best Local Similarity   50.0%;  Pred. No. 0.51;
  Matches   28;  Conservative    7;  Mismatches    8;  Indels   13;  Gaps    3;

Qy    165 RPPNAPIL------STLPETTVVRRRGRSPRRRTPSPR--RRRSQSPRRRRSQSRE 212
          :|  :| |       | |     ||| |||:||:|||| | ||:|||| ||:||:
Db    222 KPRRSPTLRYRRSRSRSP-----RRRSRSPKRRSPSPRRERHRSKSPRRHRSRSRD 272


RESULT 12
US-09-863-054-19
; Sequence 19, Application US/09863054
;   GENERAL INFORMATION:
;        APPLICANT: Chisari, Francis V.
;        TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
;                            Lymphocyte Responses to Hepatitis B Virus
;        NUMBER OF SEQUENCES: 57
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
```

```
;               COMPUTER: IBM PC compatible
;               OPERATING SYSTEM: PC-DOS/MS-DOS
;               SOFTWARE: PatentIn Release #1.0, Version #1.30
;          CURRENT APPLICATION DATA:
;               APPLICATION NUMBER: US/09/863,054
;               FILING DATE: 21-May-2001
;               CLASSIFICATION: <Unknown>
;          PRIOR APPLICATION DATA:
;               APPLICATION NUMBER: US 07/935,898
;               FILING DATE: 26-AUG-1992
;               APPLICATION NUMBER: US 08/024,120
;               FILING DATE: 26-FEB-1993
;               APPLICATION NUMBER: US 08/396,283
;               FILING DATE: 27-FEB-1995
;               APPLICATION NUMBER: US 08/463,486
;               FILING DATE: 05-JUN-1995
;          ATTORNEY/AGENT INFORMATION:
;               NAME: Lockyer, Jean M.
;               REGISTRATION NUMBER: 44,879
;               REFERENCE/DOCKET NUMBER: 014740-000421US
;          TELECOMMUNICATION INFORMATION:
;               TELEPHONE: (415) 576-0200
;               TELEFAX: (415) 576-0300
;     INFORMATION FOR SEQ ID NO: 19:
;          SEQUENCE CHARACTERISTICS:
;               LENGTH: 20 amino acids
;               TYPE: amino acid
;               STRANDEDNESS: <Unknown>
;               TOPOLOGY: linear
;          MOLECULE TYPE: peptide
;          SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-863-054-19

  Query Match             9.6%;  Score 110;  DB 5;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.085;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     132 LLWFHISCLTFGRETVIEYL 151
           ||||||||||||||||||||
Db       1 LLWFHISCLTFGRETVIEYL 20


RESULT 13
US-10-226-007-714
; Sequence 714, Application US/10226007
; GENERAL INFORMATION:
;  APPLICANT: Myriad Genetics, Inc.
;  APPLICANT:  Morham, Scott
;  APPLICANT:  Zavitz, Kenton
;  APPLICANT:  Hobden, Adrian
;  TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
;  FILE REFERENCE: 5005.01
;  CURRENT APPLICATION NUMBER: US/10/226,007
;  CURRENT FILING DATE:  2002-11-15
;  PRIOR APPLICATION NUMBER: US 60/313,883
;  PRIOR FILING DATE: 2001-08-21
;  NUMBER OF SEQ ID NOS: 1673
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 714
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-226-007-714

  Query Match             9.6%;  Score 110;  DB 6;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 0.085;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     149 EYLVSFGVWIRTPPAYRPPN 168
           ||||||||||||||| |||||
Db       1 EYLVSFGVWIRTPPPYRPPN 20


RESULT 14
US-10-226-007-646
; Sequence 646, Application US/10226007
; GENERAL INFORMATION:
;  APPLICANT: Myriad Genetics, Inc.
;  APPLICANT:  Morham, Scott
;  APPLICANT:  Zavitz, Kenton
;  APPLICANT:  Hobden, Adrian
;  TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
;  FILE REFERENCE: 5005.01
;  CURRENT APPLICATION NUMBER: US/10/226,007
;  CURRENT FILING DATE:  2002-11-15
;  PRIOR APPLICATION NUMBER: US 60/313,883
;  PRIOR FILING DATE: 2001-08-21
;  NUMBER OF SEQ ID NOS: 1673
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 646
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-226-007-646

  Query Match             9.5%;  Score 109;  DB 6;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 0.1;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     157 WIRTPPAYRPPNAPILSTLP 176
           |||||| ||||||||||||||
Db       1 WIRTPPPYRPPNAPILSTLP 20


RESULT 15
US-10-226-007-679
; Sequence 679, Application US/10226007
; GENERAL INFORMATION:
;  APPLICANT: Myriad Genetics, Inc.
;  APPLICANT:  Morham, Scott
;  APPLICANT:  Zavitz, Kenton
;  APPLICANT:  Hobden, Adrian
;  TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
;  FILE REFERENCE: 5005.01
;  CURRENT APPLICATION NUMBER: US/10/226,007
;  CURRENT FILING DATE:  2002-11-15
;  PRIOR APPLICATION NUMBER: US 60/313,883
;  PRIOR FILING DATE: 2001-08-21
;  NUMBER OF SEQ ID NOS: 1673
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 679
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-226-007-679

  Query Match             9.5%;  Score 109;  DB 6;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 0.1;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     154 FGVWIRTPPAYRPPNAPILS 173
           ||||||||| ||||||||||
Db       1 FGVWIRTPPPYRPPNAPILS 20


Search completed: February  3, 2003, 09:35:00
Job time : 17.6045 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:23:57 ; Search time 153.571 Seconds
(without alignments)
902.626 Million cell updates/sec

Title: US-09-890-752A-2
Perfect score: 1148
Sequence: 1 MPLSSIFSRIGDPTVQASKL.....RRRRSQSPRRRRSQSREPQC 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1148	100.0	215	22	US-09-890-752A-2	Sequence 2, Appli
2	1048.5	91.3	212	21	US-09-791-537-6815	Sequence 6815, Ap
3	1048.5	91.3	212	22	US-09-848-616-114	Sequence 114, App
4	1048.5	91.3	212	26	US-10-243-739-54	Sequence 54, Appl
5	1048.5	91.3	212	26	US-10-244-065-54	Sequence 54, Appl
6	1040.5	90.6	212	1	PCT-US97-15500-3	Sequence 3, Appli
7	1040.5	90.6	212	21	US-09-705-547-11	Sequence 11, Appl
8	1040.5	90.6	212	21	US-09-791-537-92352	Sequence 92352, A
9	1040.5	90.6	212	23	US-09-929-955-11	Sequence 11, Appl
10	1040.5	90.6	212	25	US-10-104-966-11	Sequence 11, Appl
11	1040.5	90.6	212	27	US-60-229-175-11	Sequence 11, Appl
12	1038.5	90.5	212	21	US-09-791-537-125025	Sequence 125025,
13	1037.5	90.4	212	21	US-09-718-095-12	Sequence 12, Appl
14	1037.5	90.4	212	21	US-09-791-537-15944	Sequence 15944, A
15	1037.5	90.4	212	21	US-09-791-537-47294	Sequence 47294, A
16	1037.5	90.4	212	21	US-09-791-537-125027	Sequence 125027,
17	1037.5	90.4	212	22	US-09-848-616-107	Sequence 107, App
18	1037.5	90.4	212	22	US-09-848-616-108	Sequence 108, App
19	1037.5	90.4	212	22	US-09-848-616-109	Sequence 109, App
20	1037.5	90.4	212	25	US-10-130-915-15	Sequence 15, Appl
21	1037.5	90.4	212	26	US-10-243-739-47	Sequence 47, Appl
22	1037.5	90.4	212	26	US-10-243-739-48	Sequence 48, Appl
23	1037.5	90.4	212	26	US-10-244-065-47	Sequence 47, Appl
24	1037.5	90.4	212	26	US-10-244-065-48	Sequence 48, Appl
25	1035.5	90.2	212	21	US-09-791-537-125030	Sequence 125030,
26	1035.5	90.2	212	21	US-09-791-537-125386	Sequence 125386,
27	1035.5	90.2	212	22	US-09-848-616-112	Sequence 112, App
28	1035.5	90.2	212	22	US-09-848-616-113	Sequence 113, App
29	1035.5	90.2	212	26	US-10-243-739-52	Sequence 52, Appl
30	1035.5	90.2	212	26	US-10-243-739-53	Sequence 53, Appl
31	1035.5	90.2	212	26	US-10-244-065-52	Sequence 52, Appl
32	1035.5	90.2	212	26	US-10-244-065-53	Sequence 53, Appl
33	1034.5	90.1	212	21	US-09-791-537-125001	Sequence 125001,
34	1034.5	90.1	212	21	US-09-791-537-125036	Sequence 125036,
35	1034.5	90.1	212	21	US-09-791-537-125388	Sequence 125388,
36	1034.5	90.1	212	22	US-09-848-616-115	Sequence 115, App
37	1034.5	90.1	212	22	US-09-848-616-116	Sequence 116, App
38	1034.5	90.1	212	22	US-09-848-616-117	Sequence 117, App
39	1034.5	90.1	212	26	US-10-243-739-49	Sequence 49, Appl
40	1034.5	90.1	212	26	US-10-243-739-55	Sequence 55, Appl
41	1034.5	90.1	212	26	US-10-243-739-56	Sequence 56, Appl
42	1034.5	90.1	212	26	US-10-243-739-57	Sequence 57, Appl
43	1034.5	90.1	212	26	US-10-244-065-49	Sequence 49, Appl
44	1034.5	90.1	212	26	US-10-244-065-55	Sequence 55, Appl
45	1034.5	90.1	212	26	US-10-244-065-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1
US-09-890-752A-2
; Sequence 2, Application US/09890752A
; GENERAL INFORMATION:
; APPLICANT: Hildt, Eberhard
; APPLICANT: Hofschneider, Peter
; TITLE OF INVENTION: Particles for Gene Therapy
; FILE REFERENCE: 107070-120 (VOS-013)
; CURRENT APPLICATION NUMBER: US/09/890,752A
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: PCT/DE00/00363
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: DE 199 04 800.2
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein comprising a HBcAg, a cell-permeability-
; OTHER INFORMATION: mediating polypeptide and heterologous binding site RGD
US-09-890-752A-2

Query Match 100.0%; Score 1148; DB 22; Length 215;
Best Local Similarity 100.0%; Pred. No. 9.1e-110;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MPLSSIFSRIGDPTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPLSSIFSRIGDPTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDT 60

Qy      61 ASALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSY 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ASALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSY 120

Qy     121 VNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 VNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTV 180

Qy     181 VRRRGRSPRRRTPSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||||||||||||||||||
Db     181 VRRRGRSPRRRTPSPRRRRSQSPRRRRSQSREPQC 215


RESULT 2
US-09-791-537-6815
; Sequence 6815, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6815
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-791-537-6815

  Query Match             91.3%;  Score 1048.5;  DB 21;  Length 212;
  Best Local Similarity   97.5%;  Pred. No. 1.8e-99;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||||||
Db     190 PSPRRRRSQSPRRRRSQSREPQC 212


RESULT 3
US-09-848-616-114
; Sequence 114, Application US/09848616
; GENERAL INFORMATION:
;  APPLICANT: Sebbel, Peter
;  APPLICANT:  Dunant, Nicolas
;  APPLICANT:  Bachmann, Martin
;  APPLICANT:  Tissot, Alain
;  APPLICANT:  Lechner, Franziska
;  TITLE OF INVENTION: Molecular Antigen Array
;  FILE REFERENCE: 1700.0180002
;  CURRENT APPLICATION NUMBER: US/09/848,616
;  CURRENT FILING DATE:  2001-05-05
;  NUMBER OF SEQ ID NOS: 186
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 114
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-848-616-114

  Query Match             91.3%;  Score 1048.5;  DB 22;  Length 212;
  Best Local Similarity   97.5%;  Pred. No. 1.8e-99;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||||||
Db     190 PSPRRRRSQSPRRRRSQSREPQC 212


RESULT 4
US-10-243-739-54
; Sequence 54, Application US/10243739
; GENERAL INFORMATION:
;  APPLICANT: Bachmann, Martin F.
;  APPLICANT:  Storni, Tazio
;  APPLICANT:  Lechner, Franziska
;  TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
;  TITLE OF INVENTION:  Immune Responses Induced by Virus Like Particles
;  FILE REFERENCE: 1700.0210001
;  CURRENT APPLICATION NUMBER: US/10/243,739
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 60/318,967
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-243-739-54

  Query Match             91.3%;  Score 1048.5;  DB 26;  Length 212;
  Best Local Similarity   97.5%;  Pred. No. 1.8e-99;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||||||
Db     190 PSPRRRRSQSPRRRRSQSREPQC 212


RESULT 5
US-10-244-065-54
; Sequence 54, Application US/10244065
```

```
; GENERAL INFORMATION:
;  APPLICANT: Bachmann, Martin F.
;  APPLICANT:  Storni, Tazio
;  APPLICANT:  Maurer, Patrick
;  APPLICANT:  Tissot, Alain
;  APPLICANT:  Schwarz, Katrin
;  APPLICANT:  Meijerink, Edwin
;  APPLICANT:  Lipowsky, Gerad
;  APPLICANT:  Pumpens, Paul
;  APPLICANT:  Cielens, Indulis
;  APPLICANT:  Renhofa, Regina
;  TITLE OF INVENTION: Packaging of Immunostimulatory Substances into Virus-like Particl
;  TITLE OF INVENTION:  Method of Preparation and Use
;  FILE REFERENCE: 1700.0220001
;  CURRENT APPLICATION NUMBER: US/10/244,065
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 60/374,145
;  PRIOR FILING DATE: 2002-04-22
;  PRIOR APPLICATION NUMBER: 60/318,994
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-244-065-54

  Query Match           91.3%;  Score 1048.5;  DB 26;  Length 212;
  Best Local Similarity   97.5%;  Pred. No. 1.8e-99;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||||||
Db     190 PSPRRRRSQSPRRRRSQSREPQC 212


RESULT 6
PCT-US97-15500-3
; Sequence 3, Application PC/TUS9715500
;  GENERAL INFORMATION:
;    APPLICANT:  The General Hospital Corporation
;    TITLE OF INVENTION:  INHIBITION OF VIRAL REPLICATION
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US97/15500
;      FILING DATE:  03-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/025,370
;      FILING DATE:  03-SEP-1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  08472/705W01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  212 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US97-15500-3

  Query Match           90.6%;  Score 1040.5;  DB 1;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 1.2e-98;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 7
US-09-705-547-11
; Sequence 11, Application US/09705547
; GENERAL INFORMATION:
;  APPLICANT:  Matti Sallberg
;  APPLICANT:  Catharina Hultgren
;  TITLE OF INVENTION:  VACCINES CONTAINING RIBAVIRIN AND
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE:  TRIPEP.023AUS
;  CURRENT APPLICATION NUMBER: US/09/705,547
;  CURRENT FILING DATE:  2000-11-03
;  PRIOR APPLICATION NUMBER:  60/229,175
;  PRIOR FILING DATE:  2000-08-29
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
;   OTHER INFORMATION: (HBcAg/HBeAg) sequence
US-09-705-547-11

  Query Match           90.6%;  Score 1040.5;  DB 21;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 1.2e-98;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
```

```
             ||||||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db        75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy       133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy       193 PSPRRRRSQSPRRRRSQSREPQC 215
             |||||||||||||||||||| ||
Db       190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 8
US-09-791-537-92352
; Sequence 92352, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 92352
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-791-537-92352

```
  Query Match             90.6%;  Score 1040.5;  DB 21;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 1.2e-98;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy        13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy        73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
             ||||||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db        75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy       133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy       193 PSPRRRRSQSPRRRRSQSREPQC 215
             |||||||||||||||||||| ||
Db       190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 9
US-09-929-955-11
; Sequence 11, Application US/09929955
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
;   OTHER INFORMATION: (HBcAg/HBeAg) sequence
US-09-929-955-11

```
  Query Match             90.6%;  Score 1040.5;  DB 23;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 1.2e-98;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy        13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy        73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
             ||||||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db        75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy       133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy       193 PSPRRRRSQSPRRRRSQSREPQC 215
             |||||||||||||||||||| ||
Db       190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 10
US-10-104-966-11
; Sequence 11, Application US/10104966
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
;   OTHER INFORMATION: (HBcAg/HBeAg) sequence
US-10-104-966-11

```
  Query Match             90.6%;  Score 1040.5;  DB 25;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 1.2e-98;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy        13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy        73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
             ||||||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db        75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy       133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy       193 PSPRRRRSQSPRRRRSQSREPQC 215
             |||||||||||||||||||| ||
```

```
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 11
US-60-229-175-11
; Sequence 11, Application US/60229175
; GENERAL INFORMATION:
;  APPLICANT:  Matti Sallberg
;  APPLICANT:  Catharina Hultgren
;  TITLE OF INVENTION:  VACCINES CONTAINING RIBAVIRIN AND
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE:  TRIPEP.023PR
;  CURRENT APPLICATION NUMBER: US/60/229,175
;  CURRENT FILING DATE:  2000-08-29
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
;   OTHER INFORMATION: (HBcAg/HBeAg) sequence
US-60-229-175-11

  Query Match             90.6%;  Score 1040.5;  DB 27;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 1.2e-98;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 12
US-09-791-537-125025
; Sequence 125025, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 125025
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-791-537-125025

  Query Match             90.5%;  Score 1038.5;  DB 21;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 1.9e-98;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRALLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQI 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||||||
Db     190 PSPRRRRSQSPRRRRSQSREPQC 212


RESULT 13
US-09-718-095-12
; Sequence 12, Application US/09718095
; GENERAL INFORMATION:
;  APPLICANT:  STUYVER Lieven
;  APPLICANT:  VAN GEYT Caroline
;  APPLICANT:  DE GENDT Sija
;  TITLE OF INVENTION:  New HBV Sequences
;  FILE REFERENCE:  2551-52
;  CURRENT APPLICATION NUMBER: US/09/718,095
;  CURRENT FILING DATE:  2000-11-22
;  PRIOR APPLICATION NUMBER:  EP99870252.6
;  PRIOR FILING DATE:  1999-12-03
;  PRIOR APPLICATION NUMBER:  US60/169,287
;  PRIOR FILING DATE:  1999-12-07
;  NUMBER OF SEQ ID NOS:  169
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-718-095-12

  Query Match             90.4%;  Score 1037.5;  DB 21;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 2.4e-98;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 14
US-09-791-537-15944
; Sequence 15944, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBEI
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15944
```

```
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-791-537-15944

  Query Match             90.4%;  Score 1037.5;  DB 21;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 2.4e-98;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 15
US-09-791-537-47294
; Sequence 47294, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 47294
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-791-537-47294

  Query Match             90.4%;  Score 1037.5;  DB 21;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 2.4e-98;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVAFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


Search completed: February  3, 2003, 09:34:08
Job time : 154.571 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:25:53 ; Search time 7.86585 Seconds
(without alignments)
551.547 Million cell updates/sec

Title: US-09-890-752A-2
Perfect score: 1148
Sequence: 1 MPLSSIFSRIGDPTVQASKL.....RRRRSQSPRRRRSQSREPQC 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1040.5	90.6	212	9	US-10-104-966-11	Sequence 11, Appl
2	1040.5	90.6	212	10	US-09-929-955-11	Sequence 11, Appl
3	977.5	85.1	212	9	US-10-209-264-4	Sequence 4, Appli
4	938.5	81.8	183	10	US-09-812-862-12	Sequence 12, Appl
5	933.5	81.3	397	10	US-09-812-862-6	Sequence 6, Appli
6	928.5	80.9	724	9	US-10-068-059-12	Sequence 12, Appl
7	928.5	80.9	746	9	US-10-068-059-6	Sequence 6, Appli
8	926.5	80.7	183	10	US-09-916-230-4	Sequence 4, Appli
9	922.5	80.4	185	9	US-10-068-059-2	Sequence 2, Appli
10	922.5	80.4	185	10	US-09-916-230-5	Sequence 5, Appli
11	920.5	80.2	289	10	US-09-812-862-8	Sequence 8, Appli
12	863.5	75.2	351	10	US-09-812-862-4	Sequence 4, Appli
13	778.5	67.8	690	9	US-10-068-059-10	Sequence 10, Appl
14	772.5	67.3	709	9	US-10-068-059-8	Sequence 8, Appli
15	771.5	67.2	170	9	US-10-068-059-4	Sequence 4, Appli
16	754	65.7	152	10	US-09-916-230-6	Sequence 6, Appli
17	736	64.1	152	10	US-09-916-230-7	Sequence 7, Appli
18	607.5	52.9	346	10	US-09-812-862-2	Sequence 2, Appli
19	157	13.7	420	10	US-09-812-862-10	Sequence 10, Appl
20	114	9.9	20	10	US-09-839-447A-107	Sequence 107, App
21	114	9.9	23	10	US-09-839-447A-111	Sequence 111, App
22	112	9.8	20	10	US-09-839-447A-104	Sequence 104, App
23	111	9.7	20	9	US-09-466-035-84	Sequence 84, Appl
24	111	9.7	20	10	US-09-839-447A-101	Sequence 101, App
25	110	9.6	20	10	US-09-839-447A-105	Sequence 105, App
26	109	9.5	20	10	US-09-839-447A-100	Sequence 100, App
27	108	9.4	20	10	US-09-839-447A-99	Sequence 99, Appl
28	108	9.4	20	10	US-09-839-447A-106	Sequence 106, App
29	105	9.1	20	10	US-09-839-447A-98	Sequence 98, Appl
30	104	9.1	20	9	US-09-466-035-85	Sequence 85, Appl
31	103	9.0	20	10	US-09-839-447A-108	Sequence 108, App
32	103	9.0	20	10	US-09-839-447A-110	Sequence 110, App
33	100	8.7	20	10	US-09-839-447A-97	Sequence 97, Appl
34	100	8.7	20	10	US-09-839-447A-109	Sequence 109, App
35	99	8.6	20	10	US-09-839-447A-96	Sequence 96, Appl
36	99	8.6	20	10	US-09-839-447A-103	Sequence 103, App
37	94.5	8.2	376	10	US-09-925-301-1399	Sequence 1399, Ap
38	92.5	8.1	20	10	US-09-839-447A-102	Sequence 102, App
39	85.5	7.4	123	10	US-09-894-018-109	Sequence 109, App
40	83.5	7.3	144	10	US-09-925-299-1217	Sequence 1217, Ap
41	83.5	7.3	219	10	US-09-894-018-113	Sequence 113, App
42	83	7.2	15	9	US-09-466-035-86	Sequence 86, Appl
43	81.5	7.1	266	10	US-09-864-761-37015	Sequence 37015, A
44	79.5	6.9	80	10	US-09-864-761-35898	Sequence 35898, A
45	79	6.9	2799	9	US-10-151-736-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-104-966-11
; Sequence 11, Application US/10104966
; Patent No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPSP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
; OTHER INFORMATION: (HBcAg/HBeAg) sequence
US-10-104-966-11

Query Match 90.6%; Score 1040.5; DB 9; Length 212;
Best Local Similarity 97.0%; Pred. No. 2.4e-95;
Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189
```

```
Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 2
US-09-929-955-11
; Sequence 11, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
;  APPLICANT: Matti Sallberg
;  APPLICANT:  Catharina Hultgren
;  TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: TRIPEP.23AUS2
;  CURRENT APPLICATION NUMBER: US/09/929,955
;  CURRENT FILING DATE:  2001-08-15
;  PRIOR APPLICATION NUMBER: 09/705,547
;  PRIOR FILING DATE: 2000-11-03
;  PRIOR APPLICATION NUMBER: 60/229,175
;  PRIOR FILING DATE: 2000-08-29
;  PRIOR APPLICATION NUMBER: 60/225,767
;  PRIOR FILING DATE: 2000-08-17
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
;   OTHER INFORMATION: (HBcAg/HBeAg) sequence
US-09-929-955-11

  Query Match             90.6%;  Score 1040.5;  DB 10;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 2.4e-95;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 3
US-10-209-264-4
; Sequence 4, Application US/10209264
; Publication No. US20030003111A1
;   GENERAL INFORMATION:
;        APPLICANT: Oon, Chong Jin
;                   Lim, Gek Keow
;                   Zhao, Yi
;                   Chen, Wei Ning
;        TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
;                            USES THEREOF
;        NUMBER OF SEQUENCES: 11
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Ladas & Parry
;             STREET: 26 West 61 Street
;             CITY: New York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10023
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/209,264
;             FILING DATE: 31-Jul-2002
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/SG98/00046
;             FILING DATE: 19-JAN-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mass, Clifford J.
;             REGISTRATION NUMBER: 30,086
;             REFERENCE/DOCKET NUMBER: U-013109-7
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 708-1800
;    INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 212 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;         SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-209-264-4

  Query Match             85.1%;  Score 977.5;  DB 9;  Length 212;
  Best Local Similarity   91.1%;  Pred. No. 3.9e-89;
  Matches  184;  Conservative    5;  Mismatches    8;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           |||||||||||||| |||||||||||: |||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWDMDIDPYKEFGASAELLSFLPSDFFPSIRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||| |||||| |||||     |||:||||||| ||||| ||:
Db      75 EHCSPHHTALRQAILCWGELMNLATWVGSNLEDP-----ASRELVVSYVNVNMGLKLRQI 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||:|||||||||||| ||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVLEYLVSFGVWIRTPTAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQ 214
           |||||||||||||||||||| |
Db     190 PSPRRRRSQSPRRRRSQSRESQ 211


RESULT 4
US-09-812-862-12
; Sequence 12, Application US/09812862
; Patent No. US20020035081A1
;   GENERAL INFORMATION:
;        APPLICANT: Wands, Jack R.
;                   Scaglioni, Pier Paolo
;                   Melegari, Margherita
;        TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
;        NUMBER OF SEQUENCES: 14
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
```

```
;             APPLICATION NUMBER: US/09/812,862
;             FILING DATE: 20-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/667,073
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/492,489
;             FILING DATE: 20-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Clark, Paul T.
;             REGISTRATION NUMBER: 30,162
;             REFERENCE/DOCKET NUMBER: 00786/282001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617/542-5070
;             TELEFAX: 617/542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 183 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-812-862-12

  Query Match             81.8%;  Score 938.5;  DB 10;  Length 183;
  Best Local Similarity   95.7%;  Pred. No. 2.2e-85;
  Matches  180;  Conservative    0;  Mismatches    3;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           ||||||||||||||||||||      ||||||||||||||||||||||||||||||||| |||
Db      61 CWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGTETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     116 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRPGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           ||||| ||
Db     176 SQSRESQC 183


RESULT 5
US-09-812-862-6
; Sequence 6, Application US/09812862
; Patent No. US20020035081A1
;   GENERAL INFORMATION:
;        APPLICANT: Wands, Jack R.
;                   Scaglioni, Pier Paolo
;                   Melegari, Margherita
;        TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
;        NUMBER OF SEQUENCES: 14
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/812,862
;             FILING DATE: 20-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/667,073
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/492,489
;             FILING DATE: 20-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Clark, Paul T.
;             REGISTRATION NUMBER: 30,162
;             REFERENCE/DOCKET NUMBER: 00786/282001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617/542-5070
;             TELEFAX: 617/542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 397 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-812-862-6

  Query Match             81.3%;  Score 933.5;  DB 10;  Length 397;
  Best Local Similarity   97.3%;  Pred. No. 1.8e-84;
  Matches  179;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           ||||||||||||||||||||      ||||||||||||||||||||||||||||||||||
Db      61 CWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSR 211
           ||||
Db     176 SQSR 179


RESULT 6
US-10-068-059-12
; Sequence 12, Application US/10068059
; Patent No. US20020155434A1
; GENERAL INFORMATION:
;  APPLICANT: Mizzen, Lee A.
;  APPLICANT:  Hongwei, Liu
;  APPLICANT:  Siegel, Marvin
;  TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
;  FILE REFERENCE: 12071-017002
;  CURRENT APPLICATION NUMBER: US/10/068,059
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: US 60/266,733
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 724
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion protein
US-10-068-059-12

  Query Match             80.9%;  Score 928.5;  DB 9;  Length 724;
  Best Local Similarity   94.2%;  Pred. No. 1.2e-83;
  Matches  179;  Conservative    2;  Mismatches    2;  Indels    7;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60
```

```
Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||     |||||||:||||||||||||||||||||||||||||
Db      61 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKFRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
           :|||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRRSQSPRR 175

Qy     206 RRSQSREPQC 215
           |||||||  ||
Db     176 RRSQSRESQC 185


RESULT 7
US-10-068-059-6
; Sequence 6, Application US/10068059
; Patent No. US20020155434A1
; GENERAL INFORMATION:
;  APPLICANT: Mizzen, Lee A.
;  APPLICANT:  Hongwei, Liu
;  APPLICANT:  Siegel, Marvin
;  TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
;  FILE REFERENCE: 12071-017002
;  CURRENT APPLICATION NUMBER: US/10/068,059
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: US 60/266,733
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 746
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion protein
US-10-068-059-6

  Query Match             80.9%;  Score 928.5;  DB 9;  Length 746;
  Best Local Similarity   94.2%;  Pred. No. 1.3e-83;
  Matches  179;  Conservative    2;  Mismatches    2;  Indels    7;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      21 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 80

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||     |||||||:||||||||||||||||||||||||||||
Db      81 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKFRQLLWFHISCLTFGRETV 135

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
           :|||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db     136 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRRSQSPRR 195

Qy     206 RRSQSREPQC 215
           |||||||  ||
Db     196 RRSQSRESQC 205


RESULT 8
US-09-916-230-4
; Sequence 4, Application US/09916230
; Patent No. US20020146422A1
; GENERAL INFORMATION:
;  APPLICANT: Bachmann, Martin F.
;  APPLICANT:  Renner, Wolfgang A.
;  TITLE OF INVENTION: Compositions for Inducing Self-Specific Anti-IgE
;  TITLE OF INVENTION:  Antibodies and Uses Thereof
;  FILE REFERENCE: 1700.0140001
;  CURRENT APPLICATION NUMBER: US/09/916,230
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: US 60/221,841
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 183
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-916-230-4

  Query Match             80.7%;  Score 926.5;  DB 10;  Length 183;
  Best Local Similarity   94.7%;  Pred. No. 3.4e-84;
  Matches  178;  Conservative    0;  Mismatches    5;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||     |||||||||||||||||||||||||||||||||||||
Db      61 CWGELMTLATWVGGNLEDP-----ISRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     116 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETCVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           ||||  ||
Db     176 SQSRGSQC 183


RESULT 9
US-10-068-059-2
; Sequence 2, Application US/10068059
; Patent No. US20020155434A1
; GENERAL INFORMATION:
;  APPLICANT: Mizzen, Lee A.
;  APPLICANT:  Hongwei, Liu
;  APPLICANT:  Siegel, Marvin
;  TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
;  FILE REFERENCE: 12071-017002
;  CURRENT APPLICATION NUMBER: US/10/068,059
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: US 60/266,733
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 185
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-10-068-059-2

  Query Match             80.4%;  Score 922.5;  DB 9;  Length 185;
  Best Local Similarity   93.7%;  Pred. No. 8.6e-84;
  Matches  178;  Conservative    2;  Mismatches    3;  Indels    7;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||     |||||||:||||||||||||||||||||||||||||
Db      61 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKIRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
           :|||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRRSQSPRR 175

Qy     206 RRSQSREPQC 215
           |||||||  ||
Db     176 RRSQSRESQC 185


RESULT 10
```

```
US-09-916-230-5
; Sequence 5, Application US/09916230
; Patent No. US20020146422A1
; GENERAL INFORMATION:
;  APPLICANT: Bachmann, Martin F.
;  APPLICANT:  Renner, Wolfgang A.
;  TITLE OF INVENTION: Compositions for Inducing Self-Specific Anti-IgE
;  TITLE OF INVENTION:  Antibodies and Uses Thereof
;  FILE REFERENCE: 1700.0140001
;  CURRENT APPLICATION NUMBER: US/09/916,230
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: US 60/221,841
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 185
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-916-230-5

  Query Match             80.4%;  Score 922.5;  DB 10;  Length 185;
  Best Local Similarity   93.7%;  Pred. No. 8.6e-84;
  Matches  178;  Conservative    2;  Mismatches    3;  Indels    7;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||     |||||||:|||||||| ||||||||||||||||||
Db      61 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKIRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
           :|||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRRSQSPRR 175

Qy     206 RRSQSREPQC 215
           |||||||  ||
Db     176 RRSQSRESQC 185


RESULT 11
US-09-812-862-8
; Sequence 8, Application US/09812862
; Patent No. US20020035081A1
;   GENERAL INFORMATION:
;        APPLICANT: Wands, Jack R.
;                   Scaglioni, Pier Paolo
;                   Melegari, Margherita
;        TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
;        NUMBER OF SEQUENCES: 14
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/812,862
;             FILING DATE: 20-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/667,073
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/492,489
;             FILING DATE: 20-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Clark, Paul T.
;             REGISTRATION NUMBER: 30,162
;             REFERENCE/DOCKET NUMBER: 00786/282001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617/542-5070
;             TELEFAX: 617/542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 289 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-812-862-8

  Query Match             80.2%;  Score 920.5;  DB 10;  Length 289;
  Best Local Similarity   96.2%;  Pred. No. 2.4e-83;
  Matches  176;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           ||||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Db      61 CWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQS 210
           | :
Db     176 SST 178


RESULT 12
US-09-812-862-4
; Sequence 4, Application US/09812862
; Patent No. US20020035081A1
;   GENERAL INFORMATION:
;        APPLICANT: Wands, Jack R.
;                   Scaglioni, Pier Paolo
;                   Melegari, Margherita
;        TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
;        NUMBER OF SEQUENCES: 14
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/812,862
;             FILING DATE: 20-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/667,073
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/492,489
;             FILING DATE: 20-JUN-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Clark, Paul T.
;             REGISTRATION NUMBER: 30,162
;             REFERENCE/DOCKET NUMBER: 00786/282001
```

```
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617/542-5070
;             TELEFAX: 617/542-8906
;             TELEX: 200154
;    INFORMATION FOR SEQ ID NO: 4:
;         SEQUENCE CHARACTERISTICS:
;             LENGTH: 351 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-812-862-4

  Query Match             75.2%;  Score 863.5;  DB 10;  Length 351;
  Best Local Similarity   87.5%;  Pred. No. 1.3e-77;
  Matches  168;  Conservative    3;  Mismatches   12;  Indels    9;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 63

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           ||||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Db      64 CWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 118

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRG----RSPRRRTPSPRRRRSQSP 203
           |||||||||||||||||||||||||||||| ||:|| |    |||||||||||||||:
Db     119 IEYLVSFGVWIRTPPAYRPPNAPILSTLPEHTVIRRGGARASRSPRRRTPSPRRRRSQNS 178

Qy     204 RRRRSQSREPQC 215
           :   |    |
Db     179 QSPTSNHSPTSC 190


RESULT 13
US-10-068-059-10
; Sequence 10, Application US/10068059
; Patent No. US20020155434A1
; GENERAL INFORMATION:
;  APPLICANT: Mizzen, Lee A.
;  APPLICANT:  Hongwei, Liu
;  APPLICANT:  Siegel, Marvin
;  TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
;  FILE REFERENCE: 12071-017002
;  CURRENT APPLICATION NUMBER: US/10/068,059
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: US 60/266,733
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 690
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion protein
US-10-068-059-10

  Query Match             67.8%;  Score 778.5;  DB 9;  Length 690;
  Best Local Similarity   94.9%;  Pred. No. 6.9e-69;
  Matches  148;  Conservative    2;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||      |||||||:|||||||||||||||||||||||||||
Db      61 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKFRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR 183
           :||||||||||||||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR 151


RESULT 14
US-10-068-059-8
; Sequence 8, Application US/10068059
; Patent No. US20020155434A1
; GENERAL INFORMATION:
;  APPLICANT: Mizzen, Lee A.
;  APPLICANT:  Hongwei, Liu
;  APPLICANT:  Siegel, Marvin
;  TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
;  FILE REFERENCE: 12071-017002
;  CURRENT APPLICATION NUMBER: US/10/068,059
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: US 60/266,733
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 709
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion protein
US-10-068-059-8

  Query Match             67.3%;  Score 772.5;  DB 9;  Length 709;
  Best Local Similarity   93.6%;  Pred. No. 2.8e-68;
  Matches  147;  Conservative    2;  Mismatches    3;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      21 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 80

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||      |||||||||||||||||||||||||||||||||||
Db      81 CWGELMTLATWVGNNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 135

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRR 184
           :||||||||||||||||||||||||||||||||||  :
Db     136 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVNAK 172


RESULT 15
US-10-068-059-4
; Sequence 4, Application US/10068059
; Patent No. US20020155434A1
; GENERAL INFORMATION:
;  APPLICANT: Mizzen, Lee A.
;  APPLICANT:  Hongwei, Liu
;  APPLICANT:  Siegel, Marvin
;  TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
;  FILE REFERENCE: 12071-017002
;  CURRENT APPLICATION NUMBER: US/10/068,059
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: US 60/266,733
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 170
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion protein
US-10-068-059-4

  Query Match             67.2%;  Score 771.5;  DB 9;  Length 170;
  Best Local Similarity   95.5%;  Pred. No. 5.9e-69;
  Matches  147;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      21 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 80

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           ||||||||||||| |||||     ||||||||||||||||||||||||||||||||||||
Db      81 CWGELMTLATWVGNNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 135

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVV 181
           :||||||||||||||||||||||||||||||||||
Db     136 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVV 169


Search completed: February  3, 2003, 09:35:29
Job time : 9.86585 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:22:17 ; Search time 13.1098 Seconds
(without alignments)
482.536 Million cell updates/sec

Title: US-09-890-752A-2
Perfect score: 1148
Sequence: 1 MPLSSIFSRIGDPTVQASKL.....RRRRSQSPRRRRSQSREPQC 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1040.5	90.6	212	3	US-08-968-747-3	Sequence 3, Appli
2	1015.5	88.5	194	3	US-08-968-747-18	Sequence 18, Appl
3	1014	88.3	199	3	US-08-968-747-21	Sequence 21, Appl
4	1000.5	87.2	193	3	US-08-968-747-2	Sequence 2, Appli
5	972	84.7	346	1	US-08-105-483-217	Sequence 217, App
6	972	84.7	346	1	US-08-709-209-217	Sequence 217, App
7	972	84.7	346	1	US-08-458-101-217	Sequence 217, App
8	954.5	83.1	211	6	5196194-13	Patent No. 5196194
9	951.5	82.9	183	4	US-09-248-588-2	Sequence 2, Appli
10	938.5	81.8	183	5	PCT-US96-10602-12	Sequence 12, Appl
11	936.5	81.6	183	3	US-08-968-747-20	Sequence 20, Appl
12	933.5	81.3	397	5	PCT-US96-10602-6	Sequence 6, Appli
13	922.5	80.4	185	1	US-07-739-642-2	Sequence 2, Appli
14	922.5	80.4	185	1	US-07-739-643-2	Sequence 2, Appli
15	922.5	80.4	185	1	US-07-739-142-2	Sequence 2, Appli
16	920.5	80.2	289	5	PCT-US96-10602-8	Sequence 8, Appli
17	919.5	80.1	185	4	US-09-248-588-6	Sequence 6, Appli
18	918.5	80.0	183	4	US-09-248-588-4	Sequence 4, Appli
19	918.5	80.0	185	1	US-07-739-642-4	Sequence 4, Appli
20	918.5	80.0	185	1	US-07-739-643-4	Sequence 4, Appli
21	918.5	80.0	185	1	US-07-739-142-4	Sequence 4, Appli
22	910.5	79.3	185	1	US-07-739-642-10	Sequence 10, Appl
23	910.5	79.3	185	1	US-07-739-643-10	Sequence 10, Appl
24	910.5	79.3	185	1	US-07-739-142-10	Sequence 10, Appl
25	907.5	79.1	185	1	US-07-739-642-8	Sequence 8, Appli
26	907.5	79.1	185	1	US-07-739-643-8	Sequence 8, Appli
27	907.5	79.1	185	1	US-07-739-142-8	Sequence 8, Appli
28	863.5	75.2	351	5	PCT-US96-10602-4	Sequence 4, Appli
29	844.5	73.6	159	4	US-08-445-585-3	Sequence 3, Appli
30	821.5	71.6	154	3	US-08-968-747-1	Sequence 1, Appli
31	821.5	71.6	155	3	US-08-968-747-17	Sequence 17, Appl
32	821.5	71.6	161	3	US-08-968-747-19	Sequence 19, Appl
33	772.5	67.3	214	1	US-07-739-642-12	Sequence 12, Appl
34	772.5	67.3	214	1	US-07-739-643-12	Sequence 12, Appl
35	772.5	67.3	214	1	US-07-739-142-12	Sequence 12, Appl
36	767.5	66.9	214	1	US-07-739-642-6	Sequence 6, Appli
37	767.5	66.9	214	1	US-07-739-643-6	Sequence 6, Appli
38	767.5	66.9	214	1	US-07-739-142-6	Sequence 6, Appli
39	757.5	66.0	217	4	US-09-248-588-9	Sequence 9, Appli
40	671	58.4	188	4	US-09-248-588-7	Sequence 7, Appli
41	607.5	52.9	346	5	PCT-US96-10602-2	Sequence 2, Appli
42	194	16.9	39	3	US-08-968-747-5	Sequence 5, Appli
43	175	15.2	305	4	US-09-248-588-11	Sequence 11, Appl
44	166	14.5	305	4	US-09-248-588-13	Sequence 13, Appl
45	157	13.7	420	5	PCT-US96-10602-10	Sequence 10, Appl

ALIGNMENTS

```
RESULT 1
US-08-968-747-3
; Sequence 3, Application US/08968747
; Patent No. 6060595
;  GENERAL INFORMATION:
;    APPLICANT:  Scaglioni et al.
;    TITLE OF INVENTION:  INHIBITION OF VIRAL REPLICATION
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/968,747
;      FILING DATE:  03-SEP-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  08472/705001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  212 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-968-747-3

  Query Match             90.6%;  Score 1040.5;  DB 3;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 2.2e-106;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||||      ||||||||||||||||||||
```

```
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 2
US-08-968-747-18
; Sequence 18, Application US/08968747
; Patent No. 6060595
;  GENERAL INFORMATION:
;    APPLICANT:  Scaglioni et al.
;    TITLE OF INVENTION:  INHIBITION OF VIRAL REPLICATION
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/968,747
;      FILING DATE:  03-SEP-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  08472/705001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  194 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-968-747-18

  Query Match             88.5%;  Score 1015.5;  DB 3;  Length 194;
  Best Local Similarity   97.0%;  Pred. No. 1.1e-103;
  Matches  192;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      18 SKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSP 77
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 SKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSP 61

Qy      78 HHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHI 137
           |||||||||||||||||||||||||||||     ||||||||||||||||||||||||||
Db      62 HHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHI 116

Qy     138 SCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRR 197
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117 SCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRR 176

Qy     198 RRSQSPRRRRSQSREPQC 215
           |||||||||||||||| ||
Db     177 RRSQSPRRRRSQSRESQC 194


RESULT 3
US-08-968-747-21
; Sequence 21, Application US/08968747
; Patent No. 6060595
;  GENERAL INFORMATION:
;    APPLICANT:  Scaglioni et al.
;    TITLE OF INVENTION:  INHIBITION OF VIRAL REPLICATION
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/968,747
;      FILING DATE:  03-SEP-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  08472/705001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  199 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-968-747-21

  Query Match             88.3%;  Score 1014;  DB 3;  Length 199;
  Best Local Similarity   96.5%;  Pred. No. 1.6e-103;
  Matches  191;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      18 SKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSP 77
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 SKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSP 61

Qy      78 HHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHI 137
           |||||||||||||||||||||||||||||     | ||||||||||||||||||||||||||
Db      62 HHTALRQAILCWGELMTLATWVGVNLEDYKDDDDKSRDLVVSYVNTNMGLKFRQLLWFHI 121

Qy     138 SCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRR 197
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 SCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRR 181

Qy     198 RRSQSPRRRRSQSREPQC 215
           |||||||||||||||| ||
Db     182 RRSQSPRRRRSQSRESQC 199


RESULT 4
US-08-968-747-2
; Sequence 2, Application US/08968747
; Patent No. 6060595
;  GENERAL INFORMATION:
;    APPLICANT:  Scaglioni et al.
;    TITLE OF INVENTION:  INHIBITION OF VIRAL REPLICATION
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
```

```
;       COUNTRY:  USA
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/968,747
;       FILING DATE:  03-SEP-1997
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fraser, Janis K.
;       REGISTRATION NUMBER:  34,819
;       REFERENCE/DOCKET NUMBER:  08472/705001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617/542-5070
;       TELEFAX:  617/542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  193 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  not relevant
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-968-747-2

  Query Match             87.2%;  Score 1000.5;  DB 3;  Length 193;
  Best Local Similarity   96.0%;  Pred. No. 4.7e-102;
  Matches  190;  Conservative    1;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      18 SKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSP 77
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 SKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSP 60

Qy      78 HHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHI 137
           |||||||||||||||||||||||||||||      ||||||||||:|||||||||||| |
Db      61 HHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVDTNMGLKFRQLLWFMI 115

Qy     138 SCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRR 197
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 SCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRR 175

Qy     198 RRSQSPRRRRSQSREPQC 215
           |||||||||||||||| ||
Db     176 RRSQSPRRRRSQSRESQC 193


RESULT 5
US-08-105-483-217
; Sequence 217, Application US/08105483
; Patent No. 5494807
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/105,483
;      FILING DATE:  12-AUG-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  217:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  346 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-105-483-217

  Query Match             84.7%;  Score 972;  DB 1;  Length 346;
  Best Local Similarity   88.7%;  Pred. No. 1.5e-98;
  Matches  189;  Conservative    2;  Mismatches    6;  Indels   16;  Gaps    2;

Qy       3 LSSIFSRIGDPTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAS 62
           :||| :| |||                |||||||||||||||||||||||||||||||||
Db     150 ISSISARTGDPVT-----------NMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAS 198

Qy      63 ALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVN 122
           |||||||||||||||||||||||||||||||||||||||||||||||||      ||||||||||
Db     199 ALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVN 253

Qy     123 TNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVR 182
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     254 TNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVR 313

Qy     183 RRGRSPRRRTPSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||||||||||||||| ||
Db     314 RRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC 346


RESULT 6
US-08-709-209-217
; Sequence 217, Application US/08709209
; Patent No. 5762938
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  462
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/709,209
;      FILING DATE:  21-AUG-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/105,483
;      FILING DATE:  12-AUG-1993
;      APPLICATION NUMBER:  US 07/847,951
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2400
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  217:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  346 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-709-209-217

  Query Match             84.7%;  Score 972;  DB 1;  Length 346;
  Best Local Similarity   88.7%;  Pred. No. 1.5e-98;
  Matches  189;  Conservative    2;  Mismatches    6;  Indels   16;  Gaps    2;

Qy       3 LSSIFSRIGDPTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAS 62
           :||| :| |||           ||||||||||||||||||||||||||||||||||||
Db     150 ISSISARTGDPVT-----------NMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAS 198

Qy      63 ALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVN 122
           ||||||||||||||||||||||||||||||||||||||||||||||     ||||||||||
Db     199 ALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVN 253

Qy     123 TNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVR 182
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     254 TNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVR 313

Qy     183 RRGRSPRRRTPSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||||||||||||| ||
Db     314 RRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC 346


RESULT 7
US-08-458-101-217
; Sequence 217, Application US/08458101
; Patent No. 5766599
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Perkus, Marion E.
;    APPLICANT:  Taylor, Jill
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  No. 5766599ton, Elizabeth K.
;    APPLICANT:  Riviere, Michel
;    APPLICANT:  de Taisne, Charles
;    APPLICANT:  Limbach, Keith J.
;    APPLICANT:  Johnson, Gerard P.
;    APPLICANT:  Pincus, Steven E.
;    APPLICANT:  Cox, William I.
;    APPLICANT:  Audonnet, Jean-Christophe Francis
;    APPLICANT:  Gettig, Russell Robert
;    TITLE OF INVENTION:  GENETICALLY ENGINEERED VACCINE
;    TITLE OF INVENTION:  STRAIN
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,101
;      FILING DATE:  01-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2740
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  217:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  346 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-458-101-217

  Query Match             84.7%;  Score 972;  DB 1;  Length 346;
  Best Local Similarity   88.7%;  Pred. No. 1.5e-98;
  Matches  189;  Conservative    2;  Mismatches    6;  Indels   16;  Gaps    2;

Qy       3 LSSIFSRIGDPTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAS 62
           :||| :| |||           ||||||||||||||||||||||||||||||||||||
Db     150 ISSISARTGDPVT-----------NMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAS 198

Qy      63 ALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVN 122
           ||||||||||||||||||||||||||||||||||||||||||||||     ||||||||||
Db     199 ALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVN 253

Qy     123 TNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVR 182
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     254 TNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVR 313

Qy     183 RRGRSPRRRTPSPRRRRSQSPRRRRSQSREPQC 215
           |||||||||||||||||||||||||||||| ||
Db     314 RRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC 346


RESULT 8
5196194-13
;Patent No. 5196194
;    APPLICANT: RUTTER, WILLIAM J.;GOODMAN, HOWARD M.
;    TITLE OF INVENTION: VACCINES CONTAINING HEPATITIS B S-PROTEIN
;    NUMBER OF SEQUENCES: 21
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/679,621
;      FILING DATE: 7-DEC-1984
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 513,055
;      FILING DATE: 12-JUL-1983
;      APPLICATION NUMBER: 107,267
;      FILING DATE: 21-DEC-1979
;      APPLICATION NUMBER: 41,909
;      FILING DATE: 24-MAY-1979
;SEQ ID NO:13:
;      LENGTH: 211
5196194-13

  Query Match             83.1%;  Score 954.5;  DB 6;  Length 211;
  Best Local Similarity   91.2%;  Pred. No. 6e-97;
  Matches  187;  Conservative    3;  Mismatches    6;  Indels    9;  Gaps    4;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           |||||||||||||| |||||||||||||||||||||||||||||||||||| ||||||||
Db      14 PTVQASKLCLGWLWG-DIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASAGYREALESP 72

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||| ||||||| |||||      |||||||: ||||:||| |||
Db      73 EHCSPHHTALRQAILCWGEL-TLATWVGNNLEDP-----ASRDLVVNVVNTNVGLKIRQL 126

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRR 190
           ||||||||||||:||||||||||||||||||||||||||||||||||||||  |||||||
Db     127 LWFHISCLTFGRETVLEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRR 186

Qy     191 RTPSPRRRRSQSPRRRRSQSREPQC 215
```

```
                  |||||||||| |||||||||| ||
Db    187 RTPSPRRRRSPSPRRRRSQSRESQC 211


RESULT 9
US-09-248-588-2
; Sequence 2, Application US/09248588
; Patent No. 6231864
; GENERAL INFORMATION:
;  APPLICANT: Birkett, Ashley J.
;  TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and
;  TITLE OF INVENTION: their Derivatives
;  FILE REFERENCE: SYN-101 4564/69529
;  CURRENT APPLICATION NUMBER: US/09/248,588
;  CURRENT FILING DATE: 1999-02-11
;  EARLIER APPLICATION NUMBER: 60/074537
;  EARLIER FILING DATE: 1998-02-12
;  NUMBER OF SEQ ID NOS: 113
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 183
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-248-588-2

  Query Match             82.9%;  Score 951.5;  DB 4;  Length 183;
  Best Local Similarity   96.8%;  Pred. No. 1.1e-96;
  Matches  182;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           ||||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Db      61 CWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           ||||| ||
Db     176 SQSRESQC 183


RESULT 10
PCT-US96-10602-12
; Sequence 12, Application PC/TUS9610602
;  GENERAL INFORMATION:
;    APPLICANT:  The General Hospital Corporation
;    TITLE OF INVENTION:  INHIBITION OF HEPATITIS B REPLICATION
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/10602
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/017,814
;      FILING DATE:  20-JUN-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/282001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  183 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US96-10602-12

  Query Match             81.8%;  Score 938.5;  DB 5;  Length 183;
  Best Local Similarity   95.7%;  Pred. No. 2.8e-95;
  Matches  180;  Conservative    0;  Mismatches    3;  Indels    5;  Gaps    1;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           ||||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Db      61 CWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGTETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     116 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRPGRSPRRRTPSPRRRRSQSPRRRR 175

Qy     208 SQSREPQC 215
           ||||| ||
Db     176 SQSRESQC 183


RESULT 11
US-08-968-747-20
; Sequence 20, Application US/08968747
; Patent No. 6060595
;  GENERAL INFORMATION:
;    APPLICANT:  Scaglioni et al.
;    TITLE OF INVENTION:  INHIBITION OF VIRAL REPLICATION
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/968,747
;      FILING DATE:  03-SEP-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  08472/705001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  183 amino acids
;      TYPE:  amino acid
```

```
;     STRANDEDNESS:  not relevant
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-08-968-747-20

  Query Match             81.6%;  Score 936.5;  DB 3;  Length 183;
  Best Local Similarity   95.7%;  Pred. No. 4.7e-95;
  Matches  180;  Conservative    1;  Mismatches    2;  Indels    5;  Gaps    1;

Qy         28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy         88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
              |||||||||||||||||||     ||||||||||:||||||||||||| |||||||||||
Db         61 CWGELMTLATWVGVNLEDP-----ASRDLVVSYVDTNMGLKFRQLLWFMISCLTFGRETV 115

Qy        148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy        208 SQSREPQC 215
              ||||| ||
Db        176 SQSRESQC 183
```

RESULT 12

```
PCT-US96-10602-6
; Sequence 6, Application PC/TUS9610602
;  GENERAL INFORMATION:
;    APPLICANT:  The General Hospital Corporation
;    TITLE OF INVENTION:  INHIBITION OF HEPATITIS B REPLICATION
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/10602
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/017,814
;      FILING DATE:  20-JUN-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/282001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  397 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US96-10602-6

  Query Match             81.3%;  Score 933.5;  DB 5;  Length 397;
  Best Local Similarity   97.3%;  Pred. No. 3e-94;
  Matches  179;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy         28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy         88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
              |||||||||||||||||||     ||||||||||||||||||||||||||||||||||||
Db         61 CWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 115

Qy        148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR 175

Qy        208 SQSR 211
              ||||
Db        176 SQSR 179
```

RESULT 13

```
US-07-739-642-2
; Sequence 2, Application US/07739642
; Patent No. 5173427
;  GENERAL INFORMATION:
;    APPLICANT:  Mallonee,, Richard L.
;    TITLE OF INVENTION:  Vectors And Hosts With Increased
;    TITLE OF INVENTION:  Expression Of HBcAg
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard R. Rodrick
;      STREET:  1 Becton Drive
;      CITY:  Franklin Lakes
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  07417-1880
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/739,642
;      FILING DATE:  19910801
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stierwalt,, Brian K.
;      REGISTRATION NUMBER:  33,213
;      REFERENCE/DOCKET NUMBER:  P-2272
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-848-5317
;      TELEFAX:  201-848-9228
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  185 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
US-07-739-642-2

  Query Match             80.4%;  Score 922.5;  DB 1;  Length 185;
  Best Local Similarity   93.7%;  Pred. No. 1.6e-93;
  Matches  178;  Conservative    2;  Mismatches    3;  Indels    7;  Gaps    2;

Qy         28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy         88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
              ||||||||||||| |||||     |||||||:||||||||| ||||||||||||||||||
Db         61 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKIRQLLWFHISCLTFGRETV 115

Qy        148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
              :|||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db        116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRRSQSPRR 175

Qy        206 RRSQSREPQC 215
```

```
                  ||||||| ||
Db    176 RRSQSRESQC 185


RESULT 14
US-07-739-643-2
; Sequence 2, Application US/07739643
; Patent No. 5175094
;  GENERAL INFORMATION:
;    APPLICANT:  Mallonee,, Richard L.
;    TITLE OF INVENTION:  Increased Expression of HBcAg
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard R. Rodrick
;      STREET:  1 Becton Drive
;      CITY:  Franklin Lakes
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  07417-1880
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/739,643
;      FILING DATE:  19910801
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stierwalt,, Brian K.
;      REGISTRATION NUMBER:  33,213
;      REFERENCE/DOCKET NUMBER:  P-2090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-848-5317
;      TELEFAX:  201-848-9228
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  185 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
US-07-739-643-2

  Query Match             80.4%;  Score 922.5;  DB 1;  Length 185;
  Best Local Similarity   93.7%;  Pred. No. 1.6e-93;
  Matches  178;  Conservative    2;  Mismatches    3;  Indels    7;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||     |||||||:|||||||| ||||||||||||||||||
Db      61 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKIRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
           :|||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRRSQSPRR 175

Qy     206 RRSQSREPQC 215
           ||||||| ||
Db     176 RRSQSRESQC 185


RESULT 15
US-07-739-142-2
; Sequence 2, Application US/07739142
; Patent No. 5175272
;  GENERAL INFORMATION:
;    APPLICANT:  Mallonee,, Richard L.
;    TITLE OF INVENTION:  DNA Sequences With Increased Expression
;    TITLE OF INVENTION:  of HBcAg
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard R. Rodrick
;      STREET:  1 Becton Drive
;      CITY:  Franklin Lakes
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  07417-1880
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/739,142
;      FILING DATE:  19910801
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stierwalt,, Brian K.
;      REGISTRATION NUMBER:  33,213
;      REFERENCE/DOCKET NUMBER:  P-2271
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-848-5317
;      TELEFAX:  201-848-9228
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  185 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
US-07-739-142-2

  Query Match             80.4%;  Score 922.5;  DB 1;  Length 185;
  Best Local Similarity   93.7%;  Pred. No. 1.6e-93;
  Matches  178;  Conservative    2;  Mismatches    3;  Indels    7;  Gaps    2;

Qy      28 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 87
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 60

Qy      88 CWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETV 147
           |||||||||||||| |||||     |||||||:|||||||| ||||||||||||||||||
Db      61 CWGELMTLATWVGNNLEDP-----ASRDLVVNYVNTNMGLKIRQLLWFHISCLTFGRETV 115

Qy     148 IEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRR--RGRSPRRRTPSPRRRRSQSPRR 205
           :|||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db     116 LEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRDRGRSPRRRTPSPRRRRSQSPRR 175

Qy     206 RRSQSREPQC 215
           ||||||| ||
Db     176 RRSQSRESQC 185


Search completed:  February  3, 2003, 09:27:10
Job time :  14.1098 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:15:41 ; Search time 34.4599 Seconds
(without alignments)
831.368 Million cell updates/sec

Title: US-09-890-752A-2
Perfect score: 1148
Sequence: 1 MPLSSIFSRIGDPTVQASKL.....RRRRSQSPRRRRSQSREPQC 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1148	100.0	215	21	AAB10597	HBV fusion protein
2	1048.5	91.3	212	23	AAU80921	Hepatitis B virus
3	1040.5	90.6	212	19	AAW50250	Hepatitis B virus
4	1040.5	90.6	212	23	AAE19793	Hepatitis B virus
5	1040.5	90.6	212	23	AAE19898	Hepatitis B virus
6	1040.5	90.6	212	23	AAE17018	Hepatitis B virus
7	1037.5	90.4	212	22	AAG66924	HBV genotype D pre
8	1037.5	90.4	212	23	AAU80914	Hepatitis B virus
9	1037.5	90.4	212	23	AAU80915	Hepatitis B virus
10	1037.5	90.4	212	23	AAU80916	Hepatitis B virus
11	1035.5	90.2	212	23	AAU80919	Hepatitis B virus
12	1035.5	90.2	212	23	AAU80920	Hepatitis B virus
13	1034.5	90.1	212	23	AAU80922	Hepatitis B virus
14	1034.5	90.1	212	23	AAU80923	Hepatitis B virus
15	1034.5	90.1	212	23	AAU80924	Hepatitis B virus
16	1031.5	89.9	212	23	AAU80928	Hepatitis B virus
17	1031.5	89.9	212	23	AAU80929	Hepatitis B virus
18	1030.5	89.8	212	23	AAU80904	Hepatitis B virus
19	1030.5	89.8	212	23	AAU80933	Hepatitis B virus
20	1030.5	89.8	212	23	AAU80934	Hepatitis B virus
21	1029.5	89.7	212	23	AAU80935	Hepatitis B virus
22	1029.5	89.7	212	23	AAU80937	Hepatitis B virus
23	1028.5	89.6	212	23	AAU80898	Hepatitis B virus
24	1028.5	89.6	212	23	AAU80910	Hepatitis B virus
25	1028.5	89.6	212	23	AAU80925	Hepatitis B virus
26	1023.5	89.2	212	23	AAU80899	Hepatitis B virus
27	1022.5	89.1	212	23	AAU80901	Hepatitis B virus
28	1020.5	88.9	212	23	AAU80903	Hepatitis B virus
29	1019.5	88.8	212	22	AAG66925	HBV genotype E pre
30	1019.5	88.8	212	23	AAU80936	Hepatitis B virus
31	1015.5	88.5	194	19	AAW50242	Hepatitis B virus
32	1015.5	88.5	212	23	AAU80902	Hepatitis B virus
33	1014.5	88.4	212	23	AAU80907	Hepatitis B virus
34	1011.5	88.1	212	23	AAU80917	Hepatitis B virus
35	1011.5	88.1	214	10	AAP90702	Deduced amino acid
36	1008.5	87.8	212	23	AAU80908	Hepatitis B virus
37	1007.5	87.8	212	22	AAG66922	HBV genotype B pre
38	1007.5	87.8	212	22	AAG66923	HBV genotype C pre
39	1007.5	87.8	212	23	AAU80931	Hepatitis B virus
40	1002.5	87.3	214	22	AAG66921	HBV genotype A pre
41	1000.5	87.2	193	19	AAW50241	Hepatitis B virus
42	999.5	87.1	198	19	AAW50252	Hepatitis B virus
43	997.5	86.9	212	21	AAY44350	Human hepatitis B
44	996.5	86.8	212	23	AAU80918	Hepatitis B virus
45	996.5	86.8	214	9	AAP80961	HBV core antigen e

ALIGNMENTS

RESULT 1
AAB10597
ID AAB10597 standard; Protein; 215 AA.
XX
AC AAB10597;
XX
DT 08-JAN-2001 (first entry)
XX
DE HBV fusion protein comprising HBcAg and RGD.
XX
KW Fusion protein; protein coat; virus-specific packaging signal; psi;
KW virus protein; cell permeability; cell-specific binding site; LHB;
KW large surface protein; core antigen; gene therapy.
XX
OS Hepatitis b virus.
OS Synthetic.
XX
PN WO200046376-A2.
XX
PD 10-AUG-2000.
XX
PF 04-FEB-2000; 2000WO-DE00363.
XX
PR 05-FEB-1999; 99DE-1004800.
XX
PA (HILD/) HILDT E.
XX
PI Hildt E, Hofschneider P;
XX
DR WPI; 2000-514959/46.
DR N-PSDB; AAA71735.
XX
PT Particle for cell-specific gene delivery, useful in gene therapy,

```
PT   comprises nucleic acid in protein coat that includes a fusion protein
PT   of viral protein, permeability peptide and cell-binding site -
XX
PS   Claim 14; Fig 2; 34pp; German.
XX
CC   This invention describes a novel particle (A), comprising a protein coat
CC   with a fusion protein (FP), and, inside the coat, a nucleic acid (I)
CC   including the sequence for a virus-specific packaging signal (psi) and a
CC   structural gene. FP contains a virus protein (VP), a peptide (P) that
CC   mediates cell permeability and a heterologous cell-specific binding site
CC   (RGD). The invention also describes (1) producing (A) in which FP
CC   contains an LHBs (large surface protein of hepatitis B virus (HBV)) and
CC   RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),
CC   (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector
CC   containing the DNA of (d). The products of the invention are used in gene
CC   therapy of cells and tissues, in vivo or ex vivo. This sequence
CC   represents a fusion protein which is described in the method of the
CC   invention.
XX
SQ   Sequence   215 AA;

  Query Match             100.0%;  Score 1148;  DB 21;  Length 215;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-109;
  Matches  215;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MPLSSIFSRIGDPTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPLSSIFSRIGDPTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDT 60

Qy      61 ASALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSY 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ASALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSY 120

Qy     121 VNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 VNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTV 180

Qy     181 VRRRGRSPRRRTPSPRRRRSQSPRRRRSQSREPQC 215
          |||||||||||||||||||||||||||||||||||
Db     181 VRRRGRSPRRRTPSPRRRRSQSPRRRRSQSREPQC 215


RESULT 2
AAU80921
ID   AAU80921 standard; Protein; 212 AA.
XX
AC   AAU80921;
XX
DT   09-APR-2002  (first entry)
XX
DE   Hepatitis B virus core antigen variant (HBcAg) #26.
XX
KW   Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KW   hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KW   measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KW   Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KW   human immunodeficiency virus infection; viral hepatitis; measles;
KW   chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
KW   cancer; chronic disease; arthritis; colitis; diabetes;
KW   multiple sclerosis.
XX
OS   Hepatitis B virus.
XX
PN   WO200185208-A2.
XX
PD   15-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-IB00741.
XX
PR   05-MAY-2000; 2000US-202341P.
XX
PA   (CYTO-) CYTOS BIOTECHNOLOGY AG.
PA   (SEBB/) SEBBEL P.
PA   (DUNA/) DUNANT N.
PA   (BACH/) BACHMANN M.
PA   (TISS/) TISSOT A.
PA   (LECH/) LECHENER F.
XX
PI   Sebbel P,  Dunant N,  Bachmann M,  Tissot A,  Lechener F;
XX
DR   WPI; 2002-055561/07.
XX
PT   New composition, useful for vaccine production, comprises antigen or
PT   antigenic determinant and non-natural molecular scaffold comprising
PT   organizer and core particle such as bacterial pilus or pilin protein
PT   -s
XX
PS   Disclosure; Page 227-228; 287pp; English.
XX
CC   The invention relates to a composition comprising: (a) a non-natural
CC   molecular scaffold (molecular scaffold) which comprises a core
CC   particle such as a bacterial pilus or pilin protein, a recombinant form
CC   of the protein, a virus-like particle or a hepatitis B virus capsid
CC   protein (HBcAg), and an organiser; and (b) an antigen or antigenic
CC   determinant, where the molecular scaffold and antigenic determinant
CC   interact to form an ordered and repetitive antigen array. Suitable
CC   antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC   protein, bee venom phospholipase, Sinbis virus E2 protein, amyloid beta
CC   derived peptides and influenza M2 antigen. The composition (or vaccine)
CC   is useful for immunisation, by administration to a subject, where the
CC   administration produces an immune response, such as humoral, cellular or
CC   protective immune response, preferably a Th type 2 T-helper (Th2)
CC   response that is specific for the antigenic determinant. The
CC   administration induces antibodies specific for the antigenic determinant
CC   of a subtype corresponding to the Th2 subtype in the subject. The subject
CC   does not generate a Th2 subtype that is specific for pilus or pilin
CC   polypeptide or antigenic determinant. The composition is useful for the
CC   production of vaccines for prevention of infectious diseases such as
CC   human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC   pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC   cancer, and chronic diseases induced or accelerated by a Th1 type
CC   immune response, such as arthritis, colitis, diabetes and multiple
CC   sclerosis. The composition is useful to generate defined self-specific
CC   antibodies and specific immune responses of the Th2 type and allows the
CC   creation of highly efficient vaccines against infectious diseases, and
CC   for treating allergy, cancer, and chronic diseases induced or accelerated
CC   by a Th1 type immune response. The present sequence is a peptide or
CC   protein incorporated into the compositions of the invention.
XX
SQ   Sequence   212 AA;

  Query Match             91.3%;  Score 1048.5;  DB 23;  Length 212;
  Best Local Similarity   97.5%;  Pred. No. 3.7e-99;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
          |||||||||||||||||||||||
Db     190 PSPRRRRSQSPRRRRSQSREPQC 212


RESULT 3
AAW50250
ID   AAW50250 standard; Protein; 212 AA.
XX
```

```
AC   AAW50250;
XX
DT   28-SEP-1998  (first entry)
XX
DE   Hepatitis B virus precore p25 polypeptide.
XX
KW   Viral replication; inhibitor; HBV; nucleocapsid; gene therapy;
KW   hepatocyte; liver; p25 protein.
XX
OS   Hepatitis B virus.
XX
PN   WO9809649-A1.
XX
PD   12-MAR-1998.
XX
PF   03-SEP-1997;   97WO-US15500.
XX
PR   03-SEP-1996;   96US-0025370.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Melegari M,  Scaglioni PP,  Wands JR;
XX
DR   WPI; 1998-193325/17.
XX
PT   DNA encoding proteins which can be incorporated with wild type
PT   nucleocapsid subunit(s) into a viral nucleocapsid - useful for
PT   inhibition of viral replication, especially hepatitis B virus
XX
PS   Claim 15; Page 35; 60pp; English.
XX
CC   This polypeptide comprises the hepatitis B virus (HBV) 25 kDa (p25)
CC   protein that is encoded by the full-length HBV precore gene.
CC   Evidence is provided that HBV replication is inhibited in the
CC   presence of high levels of HBV precore or precore-related proteins.
CC   These proteins can be incorporated into HBV nucleocapsids along
CC   with the p21 core protein (see AAW50251), which is the usual
CC   nucleocapsid component, and thereby render the nucleocapsids
CC   deficient in encapsidating HBV pregenomic RNA.  Thus, over-
CC   expression of the precore proteins, or certain variants of them,
CC   leads to transdominant inhibition of HBV replication.  Suitable
CC   inhibitory proteins include p25, p22 (see AAW50241), Met-p22 (see
CC   AAW50242), p18 (see AAW50236), Met-p18 (see AAW50237) and Met-p18-Het
CC   (see AAW50238).  The inhibitory proteins can be produced by
CC   recombinant methods using claimed expression vectors and host
CC   cells.  They can be provided exogenously to the target cells for
CC   use in inhibiting HBV replication.  Alternatively, a nucleic acid
CC   construct that directs overexpression of an inhibitory protein in
CC   target cells is used for the gene therapy of HBV infection.
XX
SQ   Sequence   212 AA;

  Query Match             90.6%;  Score 1040.5;  DB 19;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 2.5e-98;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

```
RESULT 4
AAE19793
ID   AAE19793 standard; Protein; 212 AA.
XX
AC   AAE19793;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Hepatitis B virus core antigenic protein (HBcAg).
XX
KW   Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
KW   therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
KW   polio; herpes; rabies; AIDS; foot-and-mouth disease.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..29
FT                   /label= Signal-peptide
FT   Protein         30..212
FT                   /label= Mature_HBc_protein
XX
PN   WO200177158-A1.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-GB01607.
XX
PR   07-APR-2000; 2000EP-0107118.
XX
PA   (MEDE-) MEDEVA EURO LTD.
XX
PI   Gehin A,  Gilbert R,  Stuart D,  Rowlands D;
XX
DR   WPI; 2002-239995/29.
DR   N-PSDB; AAD31509.
XX
PT   Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for
PT   the prophylactic or therapeutic treatment of humans or animals against
PT   e.g. HB virus, viral hepatitis, hepatitis C virus, influenza, or
PT   foot-and-mouth disease  -
XX
PS   Disclosure; Page 24-25; 27pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) core antigen
CC   (HBcAg) fusion proteins and polynucleotides encoding such proteins.
CC   Sequences of the invention are useful in methods of prophylactic or
CC   therapeutic vaccination or to manufacture medicaments for prophylactic
CC   or therapeutic vaccination of the human or animal body against HBV,
CC   e.g. against viral hepatitis. They are also useful as a prophylactic
CC   vaccine against e.g. hepatitis c virus, influenza, polio, herpes,
CC   rabies, acquired immune deficiency syndrome (AIDS) or foot-and-mouth
CC   disease. The present sequence is hepatitis B virus core antigenic
CC   protein (HBcAg).
XX
SQ   Sequence   212 AA;

  Query Match             90.6%;  Score 1040.5;  DB 23;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 2.5e-98;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||     |||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189
```

```
Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          ||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 5
AAE19898
ID AAE19898 standard; Protein; 212 AA.
XX
AC AAE19898;
XX
DT 18-JUN-2002 (first entry)
XX
DE Hepatitis B virus C antigen (HBcAg) and E antigen (HBeAg) sequence.
XX
KW Hepatitis B virus; HBV; infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
OS Hepatitis B virus.
XX
PN WO200213855-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB01808.
XX
PR 17-AUG-2000; 2000US-225767P.
PR 29-AUG-2000; 2000US-229175P.
PR 03-NOV-2000; 2000US-0705547.
XX
PA (TRIP-) TRIPEP AB.
XX
PI Sallberg M, Hultgren C;
XX
DR WPI; 2002-241837/29.
XX
PT Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus -
XX
PS Claim 11; Page 82; 120pp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to
CC a hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is Hepatitis B virus C antigen (HBcAg) and E antigen
CC (HBeAg) sequence.
XX
SQ Sequence 212 AA;

```
  Query Match             90.6%;  Score 1040.5;  DB 23;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 2.5e-98;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 6
AAE17018
ID AAE17018 standard; Protein; 212 AA.
XX
AC AAE17018;
XX
DT 18-APR-2002 (first entry)
XX
DE Hepatitis B virus (HBV) core antigen (HBcAg).
XX
KW Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
KW prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
KW hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
KW tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
KW dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
KW food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
KW antiprotozoal.
XX
OS Hepatitis B virus.
XX
PN WO200198333-A2.
XX
PD 27-DEC-2001.
XX
PF 22-JUN-2001; 2001WO-GB02817.
XX
PR 22-JUN-2000; 2000GB-0015308.
PR 06-OCT-2000; 2000GB-0024544.
XX
PA (CELL-) CELLTECH PHARM LTD.
XX
PI Page M, Li J, Pumpens P;
XX
DR WPI; 2002-098223/13.
DR N-PSDB; AAD27422.
XX
PT New proteins comprising a modified hepatitis B core antigen, useful as
PT a vaccine in prophylactic or therapeutic vaccination of the human or
PT animal body, particularly against hepatitis B virus infection -
XX
PS Disclosure; Page 39-40; 40pp; English.
XX
CC The invention relates to modified proteins comprising hepatitis B virus
CC (HBV) core antigen (HBcAg) wherein one or more of the four arginine
CC repeats has been deleted and the protein comprising the C-terminal
CC cysteine of HBcAg. The deleted region may be replaced by an epitope
CC from a protein other than HBcAg, in which case the HBcAg acts as a
CC carrier to present the epitope to the immune system. This chimeric
CC protein or its nucleic acid is useful as a vaccine or in a method of
CC prophylactic or therapeutic vaccination of the human or animal body,
CC particularly against HBV. The nucleic acid encoding the protein may
CC be used in gene therapy or DNA vaccination protocols. The chimeric
CC protein or its nucleic acid may also be used as the basis of a
CC prophylactic vaccine against a range of diseases, e.g. HBV, hepatitis
CC A virus (HAV), hepatitis C virus (HCV), influenza, foot-and-mouth
CC disease, polio, herpes, rabies, acquired immunodeficiency syndrome
CC (AIDS), dengue fever, yellow fever, malaria, tuberculosis, whooping
CC cough, salmonellosis, typhoid, food poisoning, diarrhoea, meningitis
CC or gonorrhea. The present sequence is Hepatitis B virus core antigen
CC (HBcAg).
XX
SQ Sequence 212 AA;

```
  Query Match             90.6%;  Score 1040.5;  DB 23;  Length 212;
  Best Local Similarity   97.0%;  Pred. No. 2.5e-98;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74
```

```
Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 7
AAG66924
ID AAG66924 standard; Protein; 212 AA.
XX
AC AAG66924;
XX
DT 19-OCT-2001 (first entry)
XX
DE HBV genotype D preCore and Core polypeptide.
XX
KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW HBsAg; antiviral; vaccine; genotype G; genotyping; HBcAg; HBeAg.
XX
OS Hepatitis B virus.
XX
PN WO200140279-A2.
XX
PD 07-JUN-2001.
XX
PF 20-NOV-2000; 2000WO-EP11526.
XX
PR 03-DEC-1999; 99EP-0870252.
PR 07-DEC-1999; 99US-0169287.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy -
XX
PS Example 3; Fig 5; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype
CC G. This genotype was found with a high prevalence in patients
CC chronically infected with HBV and residing in Europe and the USA. The
CC invention relates to a fully defined sequence of 3248 nucleotides as
CC given in specification, a sequence with 92% identity to the given
CC sequence, or sequence that is degenerate to the mentioned sequences.
CC These polynucleotides are useful for HBV genotyping. The proteins
CC encoded by the polynucleotides are useful for detecting antibodies in
CC a biological sample. Ligands that bind to the proteins and antibodies
CC directed against the proteins are useful for detecting the proteins
CC and for detecting HBcAg and HBeAg (precore precursor proteins). They
CC are also useful for preparing a vaccine or medicament for treating
CC HBV infections. The present sequence is provided in an amino acid
CC sequence alignment of the preCore and Core regions of the different
CC HBV genotypes.
XX
SQ Sequence 212 AA;

```
Query Match             90.4%;  Score 1037.5;  DB 22;  Length 212;
Best Local Similarity   96.6%;  Pred. No. 5e-98;
Matches  196;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;
```

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 8
AAU80914
ID AAU80914 standard; Protein; 212 AA.
XX
AC AAU80914;
XX
DT 09-APR-2002 (first entry)
XX
DE Hepatitis B virus core antigen variant (HBcAg) #19.
XX
KW Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KW Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KW human immunodeficiency virus infection; viral hepatitis; measles;
KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
KW cancer; chronic disease; arthritis; colitis; diabetes;
KW multiple sclerosis.
XX
OS Hepatitis B virus.
XX
PN WO200185208-A2.
XX
PD 15-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-IB00741.
XX
PR 05-MAY-2000; 2000US-202341P.
XX
PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
PA (SEBB/) SEBBEL P.
PA (DUNA/) DUNANT N.
PA (BACH/) BACHMANN M.
PA (TISS/) TISSOT A.
PA (LECH/) LECHENER F.
XX
PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechener F;
XX
DR WPI; 2002-055561/07.
XX
PT New composition, useful for vaccine production, comprises antigen or
PT antigenic determinant and non-natural molecular scaffold comprising
PT organizer and core particle such as bacterial pilus or pilin protein
PT -s
XX
PS Disclosure; Page 222; 287pp; English.
XX
CC The invention relates to a composition comprising: (a) a non-natural
CC molecular scaffold (molecular scaffold) which comprises a core
CC particle such as a bacterial pilus or pilin protein, a recombinant form
CC of the protein, a virus-like particle or a hepatitis B virus capsid
CC protein (HBcAg), and an organiser; and (b) an antigen or antigenic
CC determinant, where the molecular scaffold and antigenic determinant
CC interact to form an ordered and repetitive antigen array. Suitable
CC antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC protein, bee venom phospholipase, Sinbis virus E2 protein, amyloid beta

CC derived peptides and influenza M2 antigen. The composition (or vaccine)
CC is useful for immunisation, by administration to a subject, where the
CC administration produces an immune response, such as humoral, cellular or
CC protective immune response, preferably a Th type 2 T-helper (Th2)
CC response that is specific for the antigenic determinant. The
CC administration induces antibodies specific for the antigenic determinant
CC of a subtype corresponding to the Th2 subtype in the subject. The subject
CC does not generate a Th2 subtype that is specific for pilus or pilin
CC polypeptide or antigenic determinant. The composition is useful for the
CC production of vaccines for prevention of infectious diseases such as
CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC cancer, and chronic diseases induced or accelerated by a Th1 type
CC immune response, such as arthritis, colitis, diabetes and multiple
CC sclerosis. The composition is useful to generate defined self-specific
CC antibodies and specific immune responses of the Th2 type and allows the
CC creation of highly efficient vaccines against infectious diseases, and
CC for treating allergy, cancer, and chronic diseases induced or accelerated
CC by a Th1 type immune response. The present sequence is a peptide or
CC protein incorporated into the compositions of the invention.
XX
SQ Sequence 212 AA;

Query Match 90.4%; Score 1037.5; DB 23; Length 212;
Best Local Similarity 96.6%; Pred. No. 5e-98;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYRDALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 9
AAU80915
ID AAU80915 standard; Protein; 212 AA.
XX
AC AAU80915;
XX
DT 09-APR-2002 (first entry)
XX
DE Hepatitis B virus core antigen variant (HBcAg) #20.
XX
KW Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KW Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KW human immunodeficiency virus infection; viral hepatitis; measles;
KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
KW cancer; chronic disease; arthritis; colitis; diabetes;
KW multiple sclerosis.
XX
OS Hepatitis B virus.
XX
PN WO200185208-A2.
XX
PD 15-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-IB00741.
XX
PR 05-MAY-2000; 2000US-202341P.
XX

PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
PA (SEBB/) SEBBEL P.
PA (DUNA/) DUNANT N.
PA (BACH/) BACHMANN M.
PA (TISS/) TISSOT A.
PA (LECH/) LECHENER F.
XX
PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechener F;
XX
DR WPI; 2002-055561/07.
XX
PT New composition, useful for vaccine production, comprises antigen or
PT antigenic determinant and non-natural molecular scaffold comprising
PT organizer and core particle such as bacterial pilus or pilin protein
PT -s
XX
PS Disclosure; Page 223; 287pp; English.
XX
CC The invention relates to a composition comprising: (a) a non-natural
CC molecular scaffold (molecular scaffold) which comprises a core
CC particle such as a bacterial pilus or pilin protein, a recombinant form
CC of the protein, a virus-like particle or a hepatitis B virus capsid
CC protein (HBcAg), and an organiser; and (b) an antigen or antigenic
CC determinant, where the molecular scaffold and antigenic determinant
CC interact to form an ordered and repetitive antigen array. Suitable
CC antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC protein, bee venom phospholipase, Sinbis virus E2 protein, amyloid beta
CC derived peptides and influenza M2 antigen. The composition (or vaccine)
CC is useful for immunisation, by administration to a subject, where the
CC administration produces an immune response, such as humoral, cellular or
CC protective immune response, preferably a Th type 2 T-helper (Th2)
CC response that is specific for the antigenic determinant. The
CC administration induces antibodies specific for the antigenic determinant
CC of a subtype corresponding to the Th2 subtype in the subject. The subject
CC does not generate a Th2 subtype that is specific for pilus or pilin
CC polypeptide or antigenic determinant. The composition is useful for the
CC production of vaccines for prevention of infectious diseases such as
CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC cancer, and chronic diseases induced or accelerated by a Th1 type
CC immune response, such as arthritis, colitis, diabetes and multiple
CC sclerosis. The composition is useful to generate defined self-specific
CC antibodies and specific immune responses of the Th2 type and allows the
CC creation of highly efficient vaccines against infectious diseases, and
CC for treating allergy, cancer, and chronic diseases induced or accelerated
CC by a Th1 type immune response. The present sequence is a peptide or
CC protein incorporated into the compositions of the invention.
XX
SQ Sequence 212 AA;

Query Match 90.4%; Score 1037.5; DB 23; Length 212;
Best Local Similarity 96.6%; Pred. No. 5e-98;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||:|||||||||||||||      ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGDLMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 10
AAU80916

ID AAU80916 standard; Protein; 212 AA.
XX
AC AAU80916;
XX
DT 09-APR-2002 (first entry)
XX
DE Hepatitis B virus core antigen variant (HBcAg) #21.
XX
KW Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KW Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KW human immunodeficiency virus infection; viral hepatitis; measles;
KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
KW cancer; chronic disease; arthritis; colitis; diabetes;
KW multiple sclerosis.
XX
OS Hepatitis B virus.
XX
PN WO200185208-A2.
XX
PD 15-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-IB00741.
XX
PR 05-MAY-2000; 2000US-202341P.
XX
PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
PA (SEBB/) SEBBEL P.
PA (DUNA/) DUNANT N.
PA (BACH/) BACHMANN M.
PA (TISS/) TISSOT A.
PA (LECH/) LECHENER F.
XX
PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechener F;
XX
DR WPI; 2002-055561/07.
XX
PT New composition, useful for vaccine production, comprises antigen or
PT antigenic determinant and non-natural molecular scaffold comprising
PT organizer and core particle such as bacterial pilus or pilin protein
PT -s
XX
PS Disclosure; Page 223-224; 287pp; English.
XX
CC The invention relates to a composition comprising: (a) a non-natural
CC molecular scaffold (molecular scaffold) which comprises a core
CC particle such as a bacterial pilus or pilin protein, a recombinant form
CC of the protein, a virus-like particle or a hepatitis B virus capsid
CC protein (HBcAg), and an organiser; and (b) an antigen or antigenic
CC determinant, where the molecular scaffold and antigenic determinant
CC interact to form an ordered and repetitive antigen array. Suitable
CC antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC protein, bee venom phospholipase, Sinbis virus E2 protein, amyloid beta
CC derived peptides and influenza M2 antigen. The composition (or vaccine)
CC is useful for immunisation, by administration to a subject, where the
CC administration produces an immune response, such as humoral, cellular or
CC protective immune response, preferably a Th type 2 T-helper (Th2)
CC response that is specific for the antigenic determinant. The
CC administration induces antibodies specific for the antigenic determinant
CC of a subtype corresponding to the Th2 subtype in the subject. The subject
CC does not generate a Th2 subtype that is specific for pilus or pilin
CC polypeptide or antigenic determinant. The composition is useful for the
CC production of vaccines for prevention of infectious diseases such as
CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC cancer, and chronic diseases induced or accelerated by a Th1 type
CC immune response, such as arthritis, colitis, diabetes and multiple
CC sclerosis. The composition is useful to generate defined self-specific
CC antibodies and specific immune responses of the Th2 type and allows the
CC creation of highly efficient vaccines against infectious diseases, and
CC for treating allergy, cancer, and chronic diseases induced or accelerated
CC by a Th1 type immune response. The present sequence is a peptide or
CC protein incorporated into the compositions of the invention.
XX
SQ Sequence 212 AA;

Query Match 90.4%; Score 1037.5; DB 23; Length 212;
Best Local Similarity 96.6%; Pred. No. 5e-98;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVAFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 11
AAU80919
ID AAU80919 standard; Protein; 212 AA.
XX
AC AAU80919;
XX
DT 09-APR-2002 (first entry)
XX
DE Hepatitis B virus core antigen variant (HBcAg) #24.
XX
KW Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KW Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KW human immunodeficiency virus infection; viral hepatitis; measles;
KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
KW cancer; chronic disease; arthritis; colitis; diabetes;
KW multiple sclerosis.
XX
OS Hepatitis B virus.
XX
PN WO200185208-A2.
XX
PD 15-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-IB00741.
XX
PR 05-MAY-2000; 2000US-202341P.
XX
PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
PA (SEBB/) SEBBEL P.
PA (DUNA/) DUNANT N.
PA (BACH/) BACHMANN M.
PA (TISS/) TISSOT A.
PA (LECH/) LECHENER F.
XX
PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechener F;
XX
DR WPI; 2002-055561/07.
XX
PT New composition, useful for vaccine production, comprises antigen or
PT antigenic determinant and non-natural molecular scaffold comprising
PT organizer and core particle such as bacterial pilus or pilin protein
PT -s
XX
PS Disclosure; Page 226; 287pp; English.
XX
CC The invention relates to a composition comprising: (a) a non-natural

CC molecular scaffold (molecular scaffold) which comprises a core
CC particle such as a bacterial pilus or pilin protein, a recombinant form
CC of the protein, a virus-like particle or a hepatitis B virus capsid
CC protein (HBcAg), and an organiser; and (b) an antigen or antigenic
CC determinant, where the molecular scaffold and antigenic determinant
CC interact to form an ordered and repetitive antigen array. Suitable
CC antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC protein, bee venom phospholipase, Sinbis virus E2 protein, amyloid beta
CC derived peptides and influenza M2 antigen. The composition (or vaccine)
CC is useful for immunisation, by administration to a subject, where the
CC administration produces an immune response, such as humoral, cellular or
CC protective immune response, preferably a Th type 2 T-helper (Th2)
CC response that is specific for the antigenic determinant. The
CC administration induces antibodies specific for the antigenic determinant
CC of a subtype corresponding to the Th2 subtype in the subject. The subject
CC does not generate a Th2 subtype that is specific for pilus or pilin
CC polypeptide or antigenic determinant. The composition is useful for the
CC production of vaccines for prevention of infectious diseases such as
CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC cancer, and chronic diseases induced or accelerated by a Th1 type
CC immune response, such as arthritis, colitis, diabetes and multiple
CC sclerosis. The composition is useful to generate defined self-specific
CC antibodies and specific immune responses of the Th2 type and allows the
CC creation of highly efficient vaccines against infectious diseases, and
CC for treating allergy, cancer, and chronic diseases induced or accelerated
CC by a Th1 type immune response. The present sequence is a peptide or
CC protein incorporated into the compositions of the invention.
XX
SQ Sequence 212 AA;

Query Match 90.2%; Score 1035.5; DB 23; Length 212;
Best Local Similarity 96.6%; Pred. No. 8e-98;
Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDNASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 12
AAU80920
ID AAU80920 standard; Protein; 212 AA.
XX
AC AAU80920;
XX
DT 09-APR-2002 (first entry)
XX
DE Hepatitis B virus core antigen variant (HBcAg) #25.
XX
KW Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KW Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KW human immunodeficiency virus infection; viral hepatitis; measles;
KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
KW cancer; chronic disease; arthritis; colitis; diabetes;
KW multiple sclerosis.
XX
OS Hepatitis B virus.
XX

PN WO200185208-A2.
XX
PD 15-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-IB00741.
XX
PR 05-MAY-2000; 2000US-202341P.
XX
PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
PA (SEBB/) SEBBEL P.
PA (DUNA/) DUNANT N.
PA (BACH/) BACHMANN M.
PA (TISS/) TISSOT A.
PA (LECH/) LECHENER F.
XX
PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechener F;
XX
DR WPI; 2002-055561/07.
XX
PT New composition, useful for vaccine production, comprises antigen or
PT antigenic determinant and non-natural molecular scaffold comprising
PT organizer and core particle such as bacterial pilus or pilin protein
PT -s
XX
PS Disclosure; Page 227; 287pp; English.
XX
CC The invention relates to a composition comprising: (a) a non-natural
CC molecular scaffold (molecular scaffold) which comprises a core
CC particle such as a bacterial pilus or pilin protein, a recombinant form
CC of the protein, a virus-like particle or a hepatitis B virus capsid
CC protein (HBcAg), and an organiser; and (b) an antigen or antigenic
CC determinant, where the molecular scaffold and antigenic determinant
CC interact to form an ordered and repetitive antigen array. Suitable
CC antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC protein, bee venom phospholipase, Sinbis virus E2 protein, amyloid beta
CC derived peptides and influenza M2 antigen. The composition (or vaccine)
CC is useful for immunisation, by administration to a subject, where the
CC administration produces an immune response, such as humoral, cellular or
CC protective immune response, preferably a Th type 2 T-helper (Th2)
CC response that is specific for the antigenic determinant. The
CC administration induces antibodies specific for the antigenic determinant
CC of a subtype corresponding to the Th2 subtype in the subject. The subject
CC does not generate a Th2 subtype that is specific for pilus or pilin
CC polypeptide or antigenic determinant. The composition is useful for the
CC production of vaccines for prevention of infectious diseases such as
CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC cancer, and chronic diseases induced or accelerated by a Th1 type
CC immune response, such as arthritis, colitis, diabetes and multiple
CC sclerosis. The composition is useful to generate defined self-specific
CC antibodies and specific immune responses of the Th2 type and allows the
CC creation of highly efficient vaccines against infectious diseases, and
CC for treating allergy, cancer, and chronic diseases induced or accelerated
CC by a Th1 type immune response. The present sequence is a peptide or
CC protein incorporated into the compositions of the invention.
XX
SQ Sequence 212 AA;

Query Match 90.2%; Score 1035.5; DB 23; Length 212;
Best Local Similarity 96.6%; Pred. No. 8e-98;
Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

```
Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHICCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189
```

```
Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          |||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 13
AAU80922
ID AAU80922 standard; Protein; 212 AA.
XX
AC AAU80922;
XX
DT 09-APR-2002 (first entry)
XX
DE Hepatitis B virus core antigen variant (HBcAg) #27.
XX
KW Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KW Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KW human immunodeficiency virus infection; viral hepatitis; measles;
KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
KW cancer; chronic disease; arthritis; colitis; diabetes;
KW multiple sclerosis.
XX
OS Hepatitis B virus.
XX
PN WO200185208-A2.
XX
PD 15-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-IB00741.
XX
PR 05-MAY-2000; 2000US-202341P.
XX
PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
PA (SEBB/) SEBBEL P.
PA (DUNA/) DUNANT N.
PA (BACH/) BACHMANN M.
PA (TISS/) TISSOT A.
PA (LECH/) LECHENER F.
XX
PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechener F;
XX
DR WPI; 2002-055561/07.
XX
PT New composition, useful for vaccine production, comprises antigen or
PT antigenic determinant and non-natural molecular scaffold comprising
PT organizer and core particle such as bacterial pilus or pilin protein
PT -s
XX
PS Disclosure; Page 228-229; 287pp; English.
XX
CC The invention relates to a composition comprising: (a) a non-natural
CC molecular scaffold (molecular scaffold) which comprises a core
CC particle such as a bacterial pilus or pilin protein, a recombinant form
CC of the protein, a virus-like particle or a hepatitis B virus capsid
CC protein (HBcAg), and an organiser; and (b) an antigen or antigenic
CC determinant, where the molecular scaffold and antigenic determinant
CC interact to form an ordered and repetitive antigen array. Suitable
CC antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC protein, bee venom phospholipase, Sinbis virus E2 protein, amyloid beta
CC derived peptides and influenza M2 antigen. The composition (or vaccine)
CC is useful for immunisation, by administration to a subject, where the
CC administration produces an immune response, such as humoral, cellular or
CC protective immune response, preferably a Th type 2 T-helper (Th2)
CC response that is specific for the antigenic determinant. The
CC administration induces antibodies specific for the antigenic determinant
CC of a subtype corresponding to the Th2 subtype in the subject. The subject
CC does not generate a Th2 subtype that is specific for pilus or pilin
CC polypeptide or antigenic determinant. The composition is useful for the
CC production of vaccines for prevention of infectious diseases such as
CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC cancer, and chronic diseases induced or accelerated by a Th1 type
CC immune response, such as arthritis, colitis, diabetes and multiple
CC sclerosis. The composition is useful to generate defined self-specific
CC antibodies and specific immune responses of the Th2 type and allows the
CC creation of highly efficient vaccines against infectious diseases, and
CC for treating allergy, cancer, and chronic diseases induced or accelerated
CC by a Th1 type immune response. The present sequence is a peptide or
CC protein incorporated into the compositions of the invention.
XX
SQ Sequence 212 AA;

```
  Query Match             90.1%;  Score 1034.5;  DB 23;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 1e-97;
  Matches  196;  Conservative    0;  Mismatches    2;  Indels    5;  Gaps    1;

Qy     13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLSTASALYREALESP 74

Qy     73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
          |||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db     75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy    133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy    193 PSPRRRRSQSPRRRRSQSREPQC 215
          |||||||||||||||||||| ||
Db    190 PSPRRRRSQSPRRRRSQSRESQC 212
```

RESULT 14
AAU80923
ID AAU80923 standard; Protein; 212 AA.
XX
AC AAU80923;
XX
DT 09-APR-2002 (first entry)
XX
DE Hepatitis B virus core antigen variant (HBcAg) #28.
XX
KW Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KW Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KW human immunodeficiency virus infection; viral hepatitis; measles;
KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
KW cancer; chronic disease; arthritis; colitis; diabetes;
KW multiple sclerosis.
XX
OS Hepatitis B virus.
XX
PN WO200185208-A2.
XX
PD 15-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-IB00741.
XX
PR 05-MAY-2000; 2000US-202341P.
XX
PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
PA (SEBB/) SEBBEL P.
PA (DUNA/) DUNANT N.
PA (BACH/) BACHMANN M.
PA (TISS/) TISSOT A.
PA (LECH/) LECHENER F.
XX
PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechener F;
XX
DR WPI; 2002-055561/07.
XX

```
PT   New composition, useful for vaccine production, comprises antigen or
PT   antigenic determinant and non-natural molecular scaffold comprising
PT   organizer and core particle such as bacterial pilus or pilin protein
PT   -s
XX
PS   Disclosure; Page 229-230; 287pp; English.
XX
CC   The invention relates to a composition comprising: (a) a non-natural
CC   molecular scaffold (molecular scaffold) which comprises a core
CC   particle such as a bacterial pilus or pilin protein, a recombinant form
CC   of the protein, a virus-like particle or a hepatitis B virus capsid
CC   protein (HBcAg), and an organiser; and (b) an antigen or antigenic
CC   determinant, where the molecular scaffold and antigenic determinant
CC   interact to form an ordered and repetitive antigen array. Suitable
CC   antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC   protein, bee venom phospholipase, Sinbis virus E2 protein, amyloid beta
CC   derived peptides and influenza M2 antigen. The composition (or vaccine)
CC   is useful for immunisation, by administration to a subject, where the
CC   administration produces an immune response, such as humoral, cellular or
CC   protective immune response, preferably a Th type 2 T-helper (Th2)
CC   response that is specific for the antigenic determinant. The
CC   administration induces antibodies specific for the antigenic determinant
CC   of a subtype corresponding to the Th2 subtype in the subject. The subject
CC   does not generate a Th2 subtype that is specific for pilus or pilin
CC   polypeptide or antigenic determinant. The composition is useful for the
CC   production of vaccines for prevention of infectious diseases such as
CC   human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC   pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC   cancer, and chronic diseases induced or accelerated by a Th1 type
CC   immune response, such as arthritis, colitis, diabetes and multiple
CC   sclerosis. The composition is useful to generate defined self-specific
CC   antibodies and specific immune responses of the Th2 type and allows the
CC   creation of highly efficient vaccines against infectious diseases, and
CC   for treating allergy, cancer, and chronic diseases induced or accelerated
CC   by a Th1 type immune response. The present sequence is a peptide or
CC   protein incorporated into the compositions of the invention.
XX
SQ   Sequence   212 AA;

  Query Match             90.1%;  Score 1034.5;  DB 23;  Length 212;
  Best Local Similarity   96.6%;  Pred. No. 1e-97;
  Matches  196;  Conservative    0;  Mismatches    2;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74

Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           |||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db      75 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFRQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILLTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212


RESULT 15
AAU80924
ID   AAU80924 standard; Protein; 212 AA.
XX
AC   AAU80924;
XX
DT   09-APR-2002  (first entry)
XX
DE   Hepatitis B virus core antigen variant (HBcAg) #29.
XX
KW   Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KW   hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KW   measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KW   Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KW   human immunodeficiency virus infection; viral hepatitis; measles;
KW   chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
KW   cancer; chronic disease; arthritis; colitis; diabetes;
KW   multiple sclerosis.
XX
OS   Hepatitis B virus.
XX
PN   WO200185208-A2.
XX
PD   15-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-IB00741.
XX
PR   05-MAY-2000; 2000US-202341P.
XX
PA   (CYTO-) CYTOS BIOTECHNOLOGY AG.
PA   (SEBB/) SEBBEL P.
PA   (DUNA/) DUNANT N.
PA   (BACH/) BACHMANN M.
PA   (TISS/) TISSOT A.
PA   (LECH/) LECHENER F.
XX
PI   Sebbel P,  Dunant N,  Bachmann M,  Tissot A,  Lechener F;
XX
DR   WPI; 2002-055561/07.
XX
PT   New composition, useful for vaccine production, comprises antigen or
PT   antigenic determinant and non-natural molecular scaffold comprising
PT   organizer and core particle such as bacterial pilus or pilin protein
PT   -s
XX
PS   Disclosure; Page 230; 287pp; English.
XX
CC   The invention relates to a composition comprising: (a) a non-natural
CC   molecular scaffold (molecular scaffold) which comprises a core
CC   particle such as a bacterial pilus or pilin protein, a recombinant form
CC   of the protein, a virus-like particle or a hepatitis B virus capsid
CC   protein (HBcAg), and an organiser; and (b) an antigen or antigenic
CC   determinant, where the molecular scaffold and antigenic determinant
CC   interact to form an ordered and repetitive antigen array. Suitable
CC   antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC   protein, bee venom phospholipase, Sinbis virus E2 protein, amyloid beta
CC   derived peptides and influenza M2 antigen. The composition (or vaccine)
CC   is useful for immunisation, by administration to a subject, where the
CC   administration produces an immune response, such as humoral, cellular or
CC   protective immune response, preferably a Th type 2 T-helper (Th2)
CC   response that is specific for the antigenic determinant. The
CC   administration induces antibodies specific for the antigenic determinant
CC   of a subtype corresponding to the Th2 subtype in the subject. The subject
CC   does not generate a Th2 subtype that is specific for pilus or pilin
CC   polypeptide or antigenic determinant. The composition is useful for the
CC   production of vaccines for prevention of infectious diseases such as
CC   human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC   pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC   cancer, and chronic diseases induced or accelerated by a Th1 type
CC   immune response, such as arthritis, colitis, diabetes and multiple
CC   sclerosis. The composition is useful to generate defined self-specific
CC   antibodies and specific immune responses of the Th2 type and allows the
CC   creation of highly efficient vaccines against infectious diseases, and
CC   for treating allergy, cancer, and chronic diseases induced or accelerated
CC   by a Th1 type immune response. The present sequence is a peptide or
CC   protein incorporated into the compositions of the invention.
XX
SQ   Sequence   212 AA;

  Query Match             90.1%;  Score 1034.5;  DB 23;  Length 212;
  Best Local Similarity   96.1%;  Pred. No. 1e-97;
  Matches  195;  Conservative    2;  Mismatches    1;  Indels    5;  Gaps    1;

Qy      13 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 PTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESP 74
```

```
Qy      73 EHCSPHHTALRQAILCWGELMTLATWVGVNLEDPEFRGDASRDLVVSYVNTNMGLKFRQL 132
           ||||||||||||||||||:||||||||||||||     |||||||||||||||||:||
Db      75 EHCSPHHTALRQAILCWGDLMTLATWVGVNLEDP-----ASRDLVVSYVNTNMGLKFKQL 129

Qy     133 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 LWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT 189

Qy     193 PSPRRRRSQSPRRRRSQSREPQC 215
           ||||||||||||||||||| ||
Db     190 PSPRRRRSQSPRRRRSQSRESQC 212
```

Search completed: February 3, 2003, 09:23:51
Job time : 35.4599 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:16:26 ; Search time 0.439024 Seconds
(without alignments)
1133.686 Million cell updates/sec

Title: US-09-890-752A-20
Perfect score: 61
Sequence: 1 PLSSIFSRIGDP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61	100.0	389	1	VMSA_HPBVY	P03138 hepatitis b
2	59	96.7	389	1	VMSA_HPBVA	P24025 hepatitis b
3	59	96.7	389	1	VMSA_HPBVZ	P03139 hepatitis b
4	54	88.5	400	1	VMSA_HPBVR	P03140 hepatitis b
5	46	75.4	400	1	VMSA_HPBV4	P12934 hepatitis b
6	44	72.1	400	1	VMSA_HPBVP	Q02317 hepatitis b
7	40	65.6	64	1	BD02_HUMAN	O15263 homo sapien
8	40	65.6	389	1	VMSA_HPBVL	P12911 hepatitis b
9	38	62.3	389	1	VMSA_HPBVJ	P17398 hepatitis b
10	38	62.3	389	1	VMSA_HPBVO	P17399 hepatitis b
11	38	62.3	537	1	AREH_SCHPO	Q10269 schizosacch
12	38	62.3	1786	1	UVRA_CHLTR	O84337 chlamydia t
13	37	60.7	495	1	YQR2_CAEEL	Q09307 caenorhabdi
14	37	60.7	533	1	MVIN_SYNY3	Q55179 synechocyst
15	36	59.0	334	1	DCUP_CHLMU	Q9plh7 chlamydia m
16	36	59.0	389	1	VMSA_HPBVW	P03142 hepatitis b
17	36	59.0	400	1	VMSA_HPBV2	P03141 hepatitis b
18	36	59.0	400	1	VMSA_HPBV9	P17101 hepatitis b
19	36	59.0	593	1	MP44_MCV1	Q98224 molluscum c
20	36	59.0	695	1	EFG1_SYNY3	P28371 synechocyst
21	36	59.0	701	1	EFG_MICLU	P09952 micrococcus
22	36	59.0	1081	1	PDR6_YEAST	P32767 saccharomyc
23	35	57.4	341	1	EFG_STRRA	P29541 streptomyce
24	35	57.4	389	1	VMSA_HPBVI	P17397 hepatitis b
25	35	57.4	400	1	VMSA_HPBVT	Q05496 hepatitis b
26	35	57.4	538	1	IPGD_SHIFL	Q07566 shigella fl
27	35	57.4	701	1	EFG_MYCLE	P30767 mycobacteri
28	35	57.4	701	1	EFG_MYCTU	O53790 mycobacteri
29	35	57.4	708	1	EFG1_STRCO	P40173 streptomyce
30	35	57.4	1302	1	RRPL_BTV10	P13840 bluetongue
31	34	55.7	70	1	NUSG_LIBAS	P36262 liberibacte
32	34	55.7	225	1	NTCA_SYNY3	P33779 synechocyst
33	34	55.7	247	1	YGFF_ECOLI	P52037 escherichia
34	34	55.7	250	1	VGLL_HSV6Z	P52526 human herpe
35	34	55.7	260	1	NMA_HUMAN	Q13145 homo sapien
36	34	55.7	265	1	CB23_TOBAC	P27494 nicotiana t
37	34	55.7	268	1	CB23_HORVU	P27523 hordeum vul
38	34	55.7	270	1	CB12_LYCES	P10708 lycopersico
39	34	55.7	270	1	CB12_PETHY	P13869 petunia hyb
40	34	55.7	272	1	MURI_MYCLE	P46705 mycobacteri
41	34	55.7	283	1	LGT_HELPJ	Q9zkp6 helicobacte
42	34	55.7	284	1	LGT_HELPY	O25609 helicobacte
43	34	55.7	312	1	PHO4_YEAST	P07270 saccharomyc
44	34	55.7	344	1	CDK7_CARAU	P51953 carassius a
45	34	55.7	421	1	PHFL_DESVH	P07598 desulfovibr

ALIGNMENTS

```
RESULT 1
VMSA_HPBVY
ID   VMSA_HPBVY     STANDARD;      PRT;   389 AA.
AC   P03138;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype ayw).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10418;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81012091; PubMed=399327;
RA   Galibert F., Mandart E., Fitoussi F., Tiollais P., Charnay P.;
RT   "Nucleotide sequence of the hepatitis B virus genome (subtype ayw)
RT   cloned in E. coli.";
RL   Nature 281:646-650(1979).
RN   [2]
RP   SEQUENCE FROM N.A. (CLONE PHB320).
RX   MEDLINE=85204397; PubMed=3996597;
RA   Bichko V., Pushko P., Dreilina D., Pumpen P., Gren E.Y.;
RT   "Subtype ayw variant of hepatitis B virus. DNA primary structure
RT   analysis.";
RL   FEBS Lett. 185:208-212(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Pumpen P., Kozlovskaya T.M., Borisova G.L., Byuchko V.V.,
RA   Dishler A.V., Kalic Y.V., Pudova N.V., Gren E.Y., Krievinya V.Y.,
RA   Kukain R.A.;
RT   "Synthesis of the surface antigen of hepatitis B virus in Escherichia
RT   coli.";
RL   Dokl. Biochem. 271:246-249(1984).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01460; -; NOT_ANNOTATED_CDS.
DR   EMBL; X02496; CAA26324.1; ALT_INIT.
DR   EMBL; M12393; AAA45496.1; ALT_INIT.
DR   PIR; A03703; SAVLAH.
DR   PIR; A03704; SAVLBH.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    163
FT   CHAIN       164    389       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD     309     309       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      75      75       Q -> E (IN REF. 2).
FT   CONFLICT     147     147       A -> S (IN REF. 2).
FT   CONFLICT     150     150       L -> I (IN REF. 2).
FT   CONFLICT     288     290       MTT -> TTP (IN REF. 2 AND 3).
SQ   SEQUENCE   389 AA;  42766 MW;  6DC9E682DA694F63 CRC64;

  Query Match             100.0%;  Score 61;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 0.00052;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           ||||||||||||
Db     149 PLSSIFSRIGDP 160


RESULT 2
VMSA_HPBVA
ID   VMSA_HPBVA     STANDARD;      PRT;   389 AA.
AC   P24025;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-MAR-1992 (Rel. 21, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (strain alpha1).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10411;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90266476; PubMed=2345966;
RA   Tong S., Li J., Vitvitski L., Trepo C.;
RT   "Active hepatitis B virus replication in the presence of anti-HBe is
RT   associated with viral variants containing an inactive pre-C region.";
RL   Virology 176:596-603(1990).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M32138; AAA45502.1; -.
DR   PIR; B34773; SAVLA1.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    163
FT   CHAIN       164    389       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   389 AA;  42733 MW;  B806D0D42B223BF5 CRC64;

  Query Match             96.7%;  Score 59;  DB 1;  Length 389;
  Best Local Similarity   91.7%;  Pred. No. 0.0012;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:||||||||||
Db     149 PISSIFSRIGDP 160


RESULT 3
VMSA_HPBVZ
ID   VMSA_HPBVZ     STANDARD;      PRT;   389 AA.
AC   P03139;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-OCT-1989 (Rel. 12, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adyw).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10419;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81012115; PubMed=399329;
RA   Pasek M., Goto T., Gilbert W., Zink B., Schaller H., Mackay P.,
RA   Leadbetter G., Murray K.;
RT   "Hepatitis B virus genes and their expression in E. coli.";
RL   Nature 282:575-579(1979).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02202; AAA45487.1; ALT_INIT.
DR   PIR; A93217; SAVLAJ.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    163
FT   CHAIN       164    389       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   389 AA;  42801 MW;  1069EF6B4679A669 CRC64;

  Query Match             96.7%;  Score 59;  DB 1;  Length 389;
  Best Local Similarity   91.7%;  Pred. No. 0.0012;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:||||||||||
Db     149 PISSIFSRIGDP 160


RESULT 4
VMSA_HPBVR
ID   VMSA_HPBVR     STANDARD;      PRT;   400 AA.
AC   P03140;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adr).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=106820;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83168919; PubMed=6300776;
RA   Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
RT   "The complete nucleotide sequences of the cloned hepatitis B virus
RT   DNA; subtype adr and adw.";
RL   Nucleic Acids Res. 11:1747-1757(1983).
RN   [2]
RP   SEQUENCE OF 176-350 FROM N.A.
RX   MEDLINE=85107103; PubMed=3968537;
RA   Takeshima H., Inokoshi J., Namiki M., Shimada J., Omura S.;
RT   "Structural analysis of the gene coding for hepatitis B virus surface
RT   antigen and its product.";
RL   J. Gen. Virol. 66:195-200(1985).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; V00867; CAA24234.1; ALT_INIT.
DR   PIR; A03705; SAVLA.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP          1    174
FT   CHAIN         175    400       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD       15     15       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      123    123       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   400 AA;  43685 MW;  16DB0B15C6CB697A CRC64;

  Query Match             88.5%;  Score 54;  DB 1;  Length 400;
  Best Local Similarity   83.3%;  Pred. No. 0.011;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:|||||| |||
Db     160 PISSIFSRTGDP 171


RESULT 5
VMSA_HPBV4
ID   VMSA_HPBV4     STANDARD;      PRT;   400 AA.
AC   P12934; Q67860;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adr4).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10409;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83246570; PubMed=6306594;
RA   Fujiyama A., Miyanohara A., Nozaki C., Yoneyama T., Ohtomo N.,
RA   Matsubara K.;
RT   "Cloning and structural analyses of hepatitis B virus DNAs, subtype
RT   adr.";
RL   Nucleic Acids Res. 11:4601-4610(1983).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X01587; CAA25747.1; ALT_INIT.
DR   EMBL; X01587; CAA25743.1; ALT_INIT.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP          1    174
FT   CHAIN         175    400       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD       15     15       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      123    123       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   400 AA;  43542 MW;  01053BBCD24053F9 CRC64;

  Query Match             75.4%;  Score 46;  DB 1;  Length 400;
  Best Local Similarity   75.0%;  Pred. No. 0.37;
  Matches    9;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |:||| || |||
Db     160 PISSISSRTGDP 171


RESULT 6
VMSA_HPBVP
ID   VMSA_HPBVP     STANDARD;      PRT;   400 AA.
AC   Q02317;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adw / strain Philippino/pFDW294).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=31514;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Estacio R.C., Chavez C.C., Okamoto H., Lingao A.L., Reyes M.T.,
RA   Domingo E., Mayumi M.;
RT   "Nucleotide sequence of a hepatitis B virus genome of subtype adw
RT   isolated from a Philippino: comparison with the reported three genomes
RT   of the same subtype.";
RL   J. Gastroenterol. Hepatol. 3:215-222(1988).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M57663; AAA69680.1; -.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP          1    174
FT   CHAIN         175    400       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD       15     15       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      123    123       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   400 AA;  43740 MW;  7EBD53A004136BD2 CRC64;

  Query Match             72.1%;  Score 44;  DB 1;  Length 400;
  Best Local Similarity   81.8%;  Pred. No. 0.88;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LSSIFSRIGDP 12
           :||| ||||||
Db     161 ISSISSRIGDP 171


RESULT 7
BD02_HUMAN
ID   BD02_HUMAN     STANDARD;      PRT;    64 AA.
AC   O15263;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Beta-defensin 2 precursor (hBD-2) (Skin-antimicrobial peptide 1)
DE   (SAP1).
GN   DEFB2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=97345625; PubMed=9202117;
```

```
RA   Harder J., Bartels J.H., Christophers E., Schroeder J.-M.;
RT   "A peptide antibiotic from human skin.";
RL   Nature 387:861-861(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=99051334; PubMed=9831658;
RA   Liu L., Wang L., Jia H.P., Zhao C., Heng H.H.Q., Schutte B.C.,
RA   McCray P.B. Jr., Ganz T.;
RT   "Structure and mapping of the human beta-defensin HBD-2 gene and its
RT   expression at sites of inflammation.";
RL   Gene 222:237-244(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20072673; PubMed=10603376;
RA   Diamond G., Kaiser V., Rhodes J., Russell J.P., Bevins C.L.;
RT   "Transcriptional regulation of beta-defensin gene expression in
RT   tracheal epithelial cells.";
RL   Infect. Immun. 68:113-119(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Harder J., Schroeder J.M.;
RT   "Transcriptional regulation of the human beta-defensin-2 (hBD-2).";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SYNTHESIS OF 24-64.
RX   PubMed=12010514;
RA   Kluever E., Schulz A., Forssmann W.-G., Adermann K.;
RT   "Chemical synthesis of beta-defensins and LEAP-1/hepcidin.";
RL   J. Pept. Res. 59:241-248(2002).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).
RX   MEDLINE=20490730; PubMed=10906336;
RA   Hoover D.M., Rajashankar K.R., Blumenthal R., Puri A., Oppenheim J.J.,
RA   Chertov O., Lubkowski J.;
RT   "The structure of human beta-defensin-2 shows evidence of higher
RT   order oligomerization.";
RL   J. Biol. Chem. 275:32911-32918(2000).
RN   [7]
RP   STRUCTURE BY NMR OF 28-64.
RX   MEDLINE=21571984; PubMed=11714914;
RA   Bauer F., Schweimer K., Kluver E., Conejo-Garcia J.R., Forssmann W.G.,
RA   Rosch P., Adermann K., Sticht H.;
RT   "Structure determination of human and murine beta-defensins reveals
RT   structural conservation in the absence of significant sequence
RT   similarity.";
RL   Protein Sci. 10:2470-2479(2001).
CC   -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY (POTENTIAL).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE SKIN AND RESPIRATORY TRACT.
CC   -!- INDUCTION: BY INFLAMMATION.
CC   -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY. LAP/TAP
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z71389; CAA95992.1; -.
DR   EMBL; AF040153; AAC33549.1; -.
DR   EMBL; AF071216; AAC69554.1; -.
DR   EMBL; AJ000152; CAB65126.1; -.
DR   PDB; 1FD3; 01-NOV-00.
DR   PDB; 1FD4; 01-NOV-00.
DR   PDB; 1E4Q; 26-NOV-01.
DR   Genew; HGNC:2767; DEFB4.
DR   MIM; 602215; -.
DR   InterPro; IPR001855; Defensin_beta.
DR   InterPro; IPR001271; Defensin_mammal.
DR   Pfam; PF00711; Defensin_beta; 1.
DR   SMART; SM00048; DEFSN; 1.
KW   Antibiotic; Signal; 3D-structure.
FT   SIGNAL       1     23       POTENTIAL.
FT   CHAIN       24     64       BETA-DEFENSIN 2.
FT   DISULFID    31     60
FT   DISULFID    38     53
FT   DISULFID    43     61
SQ   SEQUENCE   64 AA;  7038 MW;  05D6454CE7ACD10E CRC64;

  Query Match             65.6%;  Score 40;  DB 1;  Length 64;
  Best Local Similarity   58.3%;  Pred. No. 0.75;
  Matches    7;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 PLSSIFSRIGDP 12
          ||  :|  ||||
Db     17 PLPGVFGGIGDP 28


RESULT 8
VMSA_HPBVL
ID   VMSA_HPBVL     STANDARD;      PRT;   389 AA.
AC   P12911;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-APR-1990 (Rel. 14, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (strain lsh / chimpanzee isolate).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10414;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88258473; PubMed=2838576;
RA   Vaudin M., Wolstenholme A.J., Tsiquaye K.N., Zuckerman A.J.,
RA   Harrison T.J.;
RT   "The complete nucleotide sequence of the genome of a hepatitis B
RT   virus isolated from a naturally infected chimpanzee.";
RL   J. Gen. Virol. 69:1383-1389(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00220; BAA00159.1; -.
DR   PIR; C28885; SAVLCP.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP       1    163
FT   CHAIN      164    389       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD     4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   389 AA;  42539 MW;  87D1F93382A10DBD CRC64;

  Query Match             65.6%;  Score 40;  DB 1;  Length 389;
  Best Local Similarity   63.6%;  Pred. No. 4.9;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 LSSIFSRIGDP 12
          :||:||  |||
Db    150 ISSVFSTTGDP 160


RESULT 9
VMSA_HPBVJ
ID   VMSA_HPBVJ     STANDARD;      PRT;   389 AA.
```

```
AC   P17398;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adw / strain Japan/pJDW233).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10413;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89010694; PubMed=3171552;
RA   Okamoto H., Tsuda F., Sakugawa H., Sastrosoewignjo R.I., Imai M.,
RA   Miyakawa Y., Mayumi M.;
RT   "Typing hepatitis B virus by homology in nucleotide sequence:
RT   comparison of surface antigen subtypes.";
RL   J. Gen. Virol. 69:2575-2583(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00329; -; NOT_ANNOTATED_CDS.
DR   PIR; G28925; SAVLJ1.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    163
FT   CHAIN       164    389       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   389 AA;  42603 MW;  72F8D481924DC104 CRC64;

  Query Match             62.3%;  Score 38;  DB 1;  Length 389;
  Best Local Similarity   63.6%;  Pred. No. 12;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LSSIFSRIGDP 12
           :||| |: |||
Db     150 ISSILSKTGDP 160


RESULT 10
VMSA_HPBVO
ID   VMSA_HPBVO     STANDARD;      PRT;   389 AA.
AC   P17399;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   Major surface antigen precursor.
GN   S.
OS   Hepatitis B virus (subtype adw / strain Okinawa/pODW282).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10415;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89010694; PubMed=3171552;
RA   Okamoto H., Tsuda F., Sakugawa H., Sastrosoewignjo R.I., Imai M.,
RA   Miyakawa Y., Mayumi M.;
RT   "Typing hepatitis B virus by homology in nucleotide sequence:
RT   comparison of surface antigen subtypes.";
RL   J. Gen. Virol. 69:2575-2583(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00330; -; NOT_ANNOTATED_CDS.
DR   PIR; H28925; SAVLJ2.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
FT   PROPEP        1    163
FT   CHAIN       164    389       MAJOR SURFACE ANTIGEN.
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   389 AA;  42408 MW;  CECAC0DDA3D84A10 CRC64;

  Query Match             62.3%;  Score 38;  DB 1;  Length 389;
  Best Local Similarity   63.6%;  Pred. No. 12;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LSSIFSRIGDP 12
           :||| |: |||
Db     150 ISSILSKTGDP 160


RESULT 11
AREH_SCHPO
ID   AREH_SCHPO     STANDARD;      PRT;   537 AA.
AC   Q10269;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable sterol O-acyltransferase (EC 2.3.1.26) (Sterol-ester
DE   synthase).
GN   SPAC13G7.05.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- CATALYTIC ACTIVITY: Acyl-CoA + cholesterol = CoA + cholesterol
CC       ester.
```

```
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC       reticulum (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE STEROL O-ACYLTRANSFERASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z69729; CAA93593.1; -.
DR   InterPro; IPR002688; ACAT.
DR   Pfam; PF01800; ACAT; 1.
KW   Hypothetical protein; Transferase; Transmembrane;
KW   Endoplasmic reticulum; Acyltransferase.
FT   TRANSMEM     98    118       POTENTIAL.
FT   TRANSMEM    140    160       POTENTIAL.
FT   TRANSMEM    174    194       POTENTIAL.
FT   TRANSMEM    199    219       POTENTIAL.
FT   TRANSMEM    305    325       POTENTIAL.
FT   TRANSMEM    344    364       POTENTIAL.
FT   TRANSMEM    384    404       POTENTIAL.
FT   TRANSMEM    462    482       POTENTIAL.
FT   TRANSMEM    488    508       POTENTIAL.
FT   TRANSMEM    517    537       POTENTIAL.
FT   CARBOHYD    250    250       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   537 AA;  63098 MW;  41DBA79402B67A28 CRC64;

  Query Match             62.3%;  Score 38;  DB 1;  Length 537;
  Best Local Similarity   58.3%;  Pred. No. 16;
  Matches    7;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |  ||| |: ||
Db      79 PRKSIFDRVTDP 90


RESULT 12
UVRA_CHLTR
ID   UVRA_CHLTR     STANDARD;      PRT;  1786 AA.
AC   O84337;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Excinuclease ABC subunit A.
GN   UVRA OR CT333.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/Cx;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
CC   -!- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC       CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC       PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE
CC       AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC       STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
CC   -!- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRA, UVRB AND UVRC.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.
CC       CONTAINS FOUR ABC DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001306; AAC67928.1; -.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR004602; UvrA.
DR   Pfam; PF00005; ABC_tran; 2.
DR   ProDom; PD000006; ABC_transportr; 1.
DR   TIGRFAMs; TIGR00630; uvra; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; FALSE_NEG.
KW   SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
KW   DNA-binding; Zinc-finger; Complete proteome.
FT   NP_BIND      32     39       ATP (POTENTIAL).
FT   NP_BIND     625    632       ATP (POTENTIAL).
FT   NP_BIND     964    971       ATP (POTENTIAL).
FT   NP_BIND    1516   1523       ATP (POTENTIAL).
FT   ZN_FING     719    742       C4-TYPE.
FT   ZN_FING    1602   1628       C4-TYPE.
SQ   SEQUENCE   1786 AA;  196948 MW;  02D6862EE15DE070 CRC64;

  Query Match             62.3%;  Score 38;  DB 1;  Length 1786;
  Best Local Similarity   63.6%;  Pred. No. 57;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LSSIFSRIGDP 12
           | |:|:||| |
Db    1042 LESLFTRIGHP 1052


RESULT 13
YQR2_CAEEL
ID   YQR2_CAEEL     STANDARD;      PRT;  495 AA.
AC   Q09307;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Hypothetical 45.1 kDa protein F19C6.2 in chromosome X.
GN   F19C6.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Harris B.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z48006; CAA88050.1; -.
DR   WormPep; F19C6.2; CE03231.
KW   Hypothetical protein.
SQ   SEQUENCE   495 AA;  55814 MW;  99E9B52032B0B48C CRC64;

  Query Match             60.7%;  Score 37;  DB 1;  Length 495;
  Best Local Similarity   45.5%;  Pred. No. 23;
  Matches    5;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGD 11
           |:| :| ::||
Db     278 PVSGLFQKVGD 288
```

```
RESULT 14
MVIN_SYNY3
ID   MVIN_SYNY3     STANDARD;      PRT;   533 AA.
AC   Q55179;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Virulence factor mviN homolog.
GN   MVIN OR SLR0488.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE MVIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D64001; BAA10319.1; -.
DR   InterPro; IPR004268; MVIN_like.
DR   Pfam; PF03023; MVIN; 1.
KW   Transmembrane; Complete proteome.
FT   TRANSMEM     11     31       POTENTIAL.
FT   TRANSMEM     46     66       POTENTIAL.
FT   TRANSMEM     96    116       POTENTIAL.
FT   TRANSMEM    135    155       POTENTIAL.
FT   TRANSMEM    166    186       POTENTIAL.
FT   TRANSMEM    196    216       POTENTIAL.
FT   TRANSMEM    253    273       POTENTIAL.
FT   TRANSMEM    284    304       POTENTIAL.
FT   TRANSMEM    330    350       POTENTIAL.
FT   TRANSMEM    360    380       POTENTIAL.
FT   TRANSMEM    400    420       POTENTIAL.
FT   TRANSMEM    422    442       POTENTIAL.
FT   TRANSMEM    452    472       POTENTIAL.
FT   TRANSMEM    493    513       POTENTIAL.
SQ   SEQUENCE   533 AA;  57053 MW;  00A5B612F744C1D2 CRC64;

  Query Match             60.7%;  Score 37;  DB 1;  Length 533;
  Best Local Similarity   50.0%;  Pred. No. 25;
  Matches    6;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 PLSSIFSRIGDP 12
           |   :|||: ||
Db     300 PFLPVFSRLADP 311


RESULT 15
DCUP_CHLMU
ID   DCUP_CHLMU     STANDARD;      PRT;   334 AA.
AC   Q9PLH7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD).
GN   HEME OR TC0123.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MoPn / Nigg;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
CC   -!- CATALYTIC ACTIVITY: Uroporphyrinogen-III = coproporphyrinogen + 4
CC       CO(2).
CC   -!- PATHWAY: Porphyrin biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE UROPORPHYRINOGEN DECARBOXYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE002279; AAF39001.1; -.
DR   HSSP; P06132; 1URO.
DR   TIGR; TC0123; -.
DR   InterPro; IPR000257; Uro_decarbxyls.
DR   Pfam; PF01208; URO-D; 1.
DR   PROSITE; PS00906; UROD_1; 1.
DR   PROSITE; PS00907; UROD_2; 1.
KW   Lyase; Decarboxylase; Porphyrin biosynthesis; Complete proteome.
SQ   SEQUENCE   334 AA;  37380 MW;  199B6A88AC1F4D2D CRC64;

  Query Match             59.0%;  Score 36;  DB 1;  Length 334;
  Best Local Similarity   54.5%;  Pred. No. 24;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LSSIFSRIGDP 12
           || |: ::|||
Db     258 LSEIYRQLGDP 268


Search completed: February 3, 2003, 09:24:21
Job time : 1.43902 secs
```